

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:52:20 ; Search time 51 Seconds
(without alignments)
857.670 Million cell updates/sec

Title: US-09-825-751A-20
Perfect score: 2447
Sequence: 1 MDTPRVLLSAVELISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1798	73.5	436	2 B55452	cartilage-derived
2	927.5	37.9	501	2 JC2347	growth/differentia
3	919.5	37.6	495	2 S43294	bone morphogenetic
4	918.5	37.5	501	2 A53452	cartilage-derived
5	691	28.2	125	2 S43295	bone morphogenetic
6	549	22.4	151	2 S43296	bone morphogenetic
7	470	19.2	398	2 JH0687	bone morphogenetic
8	464	19.0	398	2 JH0688	bone morphogenetic
9	453	18.5	405	2 I50608	bone morphogenetic
10	441.5	18.0	408	1 BMHU4	bone morphogenetic
11	436.5	17.8	420	2 I49541	bone morphogenetic
12	434	17.7	393	2 S37073	bone morphogenetic
13	432.5	17.7	408	2 S38343	bone morphogenetic
14	431.5	17.6	396	1 BMHU2	bone morphogenetic
15	429.5	17.6	461	2 S52408	SPDVR1 protein - s
16	426.5	17.4	408	2 JH0801	bone morphogenetic
17	423	17.3	394	2 S45355	bone morphogenetic
18	419.5	17.1	408	2 S58791	bone morphogenetic
19	418.5	17.1	452	2 I49542	bone morphogenetic
20	417.5	17.1	353	2 I50607	bone morphogenetic
21	416.5	17.0	402	2 A45056	osteoecenic protein
22	416.5	17.0	454	1 BMHU5	bone morphogenetic
23	413.5	16.9	401	2 JH0689	bone morphogenetic
24	408.5	16.7	400	2 A49147	bone morphogenetic
25	407	16.6	426	2 JH0690	bone morphogenetic
26	401.5	16.4	431	1 BMHU7	bone morphogenetic
27	401	16.4	430	2 JQ1184	osteoecenic protein
28	395.5	16.2	588	2 A26158	decapentaplegic pr
29	394.5	16.1	510	2 A54798	Vg-1-related prote

30	394.5	16.1	513	1 BMHU6	bone morphogenetic
31	391	16.0	313	2 I51284	bone morphogenetic
32	377	15.4	354	2 S29718	gene nodal protein
33	372	15.2	427	2 A40735	TGF beta homolog d
34	351	14.3	476	2 JC4646	bone morphogenetic
35	350.5	14.3	478	2 JC4838	bone morphogenetic
36	343	14.0	365	2 T43286	cet-1 protein - Ca
37	343	14.0	455	2 A43918	TGF-beta-related p
38	334	13.6	207	2 S37618	vgr protein - rat
39	325.5	13.3	472	1 BMHU3	bone morphogenetic
40	324.5	13.3	360	2 A29619	Vg1 embryonic grow
41	319.5	13.1	372	2 C39564	Gdf-1 embryonic gr
42	318	13.0	426	1 B24248	inhibin beta-A cha
43	313.5	12.8	425	2 I47072	inhibin beta-A cha
44	310	12.7	424	1 WFPGBA	inhibin beta-A cha
45	306.5	12.5	425	1 S50898	inhibin beta-A cha

ALIGNMENTS

RESULT 1

B55452
cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 26-May-2000
C:Accession: B55452
R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko
J. Biol. Chem. 269, 28227-28234, 1994
A:Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr
A:Reference number: A55452; MUID:95050604; PMID:7961761
A:Accession: B55452
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-436 <CHA>
A:Cross-references: GB:U13661; NID:9632489; PIDN:AAA61416.1; PID:9632490
C:Superfamily: inhibin

Query Match 73.5%; Score 1798; DB 2; Length 436;
Best Local Similarity 80.8%; Pred. No. 8.1e-120;
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY	33	SSAELGSGTGMRSRKEGKQMRAPRDSADGRE---GGEPPRQDEPR---AQOPRAQEP	86
DB	2	ASAELGSAKGMRTKEGRMPRAPRENATAREPLDRQEPPEPPQEEPPQRRPPQPEAREPP	61
QY	87	GRGPRVVPHEYMLSIYRTYSIAEKLGINASFFOSSKSANTITSFVDRGLDLSHTPLRRQ	146
DB	62	GRGPRVLPHEYMLSIYRTYSIAEKLGINASFFOSSKSANTITSFVDRGLDLSHTPLRRQ	121
QY	147	KYLFVDSMLSDKEELVGAELRLFRQAPSPMPGAPLHVQLPPLSLLDARTLDPOG	206
DB	122	KYLFVDSMTLSDKEELVGAELRLFRQAPSPMPGAPLHVQLPPLSLLDARTLDPOG	179
QY	207	APPAGWEFVDVWGLRHQPKQLCLLELRAAW-GELDAGEAEARAGPQQPPPPDLRLSLGF	265
DB	180	APRGWEFVDVWGLRHQPKQLCLLELRAAWGCEGPAEADEARTPGPQQPPPPDLRLSLGF	239
QY	266	GRVRPQERALLVVTTRQKMLFAEMRQLGSA-EAAGPGAGAGGSW-----PPPS	317
DB	240	GRVTRPQERALLVVTTRQKMLFAEMRQLGSAEATVVGPGGAGSGSPPPPPPPPPPS	299
QY	318	GAPDAPWLPSPGRRRRRTAFASRHGKSRKSLKPLHVNFKELGWDWIATPLE	377
DB	300	GTPDAGLWSPSPG-RRRRTAFASRHGKSRKSLKPLHVNFKELGWDWIATPLE	358
QY	378	YEAYHCEGVCDFFLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLPTISILYIDAGN	437
DB	359	YEAYHCEGVCDFFLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLPTISILYIDAGN	418
QY	438	NVYKQYEDMVVESCGR 455	
DB	419	NVYNEVEEMVESCGR 436	

RESULT 2

JC2347
growth/differentiation factor 5 - human
C:Species: Homo sapiens (man)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Mar-2000
C:Accession: JC2347
R:Hoetten, G.; Neidhardt, H.; Jacobowsky, B.; Pohl, J.
Blochem. Biophys. Res. Commun. 204, 646-652, 1994
A>Title: Cloning and expression of recombinant human growth/differentiation factor 5.
A:Reference number: JC2347; MUID:95071375; PMID:7980526
A:Accession: JC2347
A:Molecule type: DNA
A:Residues: 1-501 <HOE>
A:Cross-references: GB:X80915; NID:9671524; PIDN:CAA56874.1; PID:9671525
C:Genetics:
A:Gene: GDB:BMP9
A:Cross-references: GDB:433948
A:Introns: 211/1
C:Superfamily: inhibin
C:Keywords: glycoprotein
F:189/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:381-382/Cleavage site: Arg-Ala (unidentified proteinase) #status predicted

Query Match 37.9%; Score 927.5; DB 2; Length 501;
Best Local Similarity 44.7%; Pred. No. 3 6e-58;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;
QY 38 GSTKGMRSRKEGKQMRAPRDSAGREGQPPQPRQDEPRAQ----- 79
DB 81 GOTGGLTPQKKDEPKLP-----PRPG-GPEPKPGHPQTRQATARTVTPKQLPGGKAP 134
QY 80 -----PRAQEPGPGPRVVPVPHYEMLSYRIYSIAEKLGINASFTQ 119
DB 135 PKAGSVSPFLKKAREGPPPREKPEFPPTTPHEYMLSYRTLSADRGKGNSSVKL 194
QY 120 SSKSANTITSFVDRGLDLSHTPLRQKYLFDVSMLSDEELVGAEELRFRQAPS----- 174
DB 195 EAGLANTITSFIDKGDDRGVP-VKQRYVFDISAL-EKDGLLGAELRLKPKSPDTAKP 252
QY 175 -APWGPAGPLHVQLPFCUS-----PLLLDARTLDPOGAPAGWEFVWQGLRH-OPWQK 228
DB 253 AAPGGGAAQL--KLSSCPSGRQASLLDVRV--PGLDGSWEVFDIWKLFNFKNSAQ 308
QY 229 LCLELRAANGELDAGAEARAGCPQPPDLRLSLGFGRRVRPPQERALLVFTSRQKN 288
DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRAARQVHEKALFLVFGRTKKRD 352
QY 289 L-FAEMREQLGSAEAGPGAGAGSWPPSGAPDARPLPSPGRRRRRTAFASRHGKRHG 347
DB 353 LFENEIKARSGQDDKT-----VVEYLFESQ-RRKRRAPLATROGKRPS 393
QY 348 KKSRLRCSKKPLHVNFKELGWDWIITAPLEYEAYHCEGVCDPPLRSHLEPTNHAIQTILM 407
DB 394 KNLKARCSKALHVNFKDMGWDDWIITAPLEYEAFHCEGLCEPLRSHLEPTNHAIQTILM 453
QY 408 NSMDPGSTPPSCVPTKLTPIISILYIDAGNNVYKQYEDMVVYESGCCR 455
DB 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVYKQYEDMVVYESGCCR 501

RESULT 3

S43294
bone morphogenetic protein-related protein (GDF5) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 17-Mar-2000
C:Accession: S43294
R:Storm, E.E.; Huynh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.
Nature 368, 639-643, 1994
A>Title: Limb alterations in brachypodism mice due to mutations in a new member of the T
A:Reference number: S43294; MUID:94195427; PMID:8145850
A:Accession: S43294

A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-495 <STO>
A:Cross-references: GB:U08337; NID:9488461; PIDN:AAA18778.1; PID:9488462
C:Superfamily: inhibin

Query Match 37.6%; Score 919.5; DB 2; Length 495;
Best Local Similarity 43.5%; Pred. No. 1.3e-57;
Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PQFQASISSSSCAELGSTKMRKQMRAPRDSAGREGQPPQPRQDE----- 74
DB 60 PGHHYGVGATNARAK-GSSGQTQAKKDEPRKMPRSG-----GSETKGPPSSQTRQAAA 113
QY 75 -----PRAQOP-----RAQEP-----PGRGPRVVPVPHYEMLSIYTT 104
DB 114 RTVTPKGQLPGGKASKAGSAPSSFLKKTRFPGTPREPKEFPPTTPHEYMLSLYRT 173
QY 105 YSTAELKGINASFFQSSKSANTITSFVDRGLDLSHTPLRQKYLFDVSMLSDEELVGA 164
DB 174 LSDADRKGGNSSVKLEAGLANITITFIDKGQDDRG-PAYRKQRYVFDISAL-EKDGLLGA 231
QY 165 ELRLFRQAP---SAPWGPAGPL-HVQLPFCUS-----PLLLDARTLDPOGAPAGWEVFD 216
DB 232 ELRLRKKPLDVAKPAVPSSGRVAQLKLSSCPSGRQPAALLDVRV--PGLDGSWEVFD 289
QY 217 VHQGLRH-OPWQGLRH-OPWQGLRH-OPWQGLRH-OPWQGLRH-OPWQGLRH-OPWQGLRH- 275
DB 290 IWKLFNFKNSAQLCLELE-AW---ERGRA-----VDLRGLGFTERTARQVHEK 333
QY 276 ALLVVFTRSORKNL-FAEMREQLGSAEAGPGAGAGSWPPSGAPDARPLPSPGRRRR 334
DB 334 ALFLVFGRTKKRDLFENEIKARSGQDDKT-----VVEYLFESQ-RRKR 374
QY 335 RTAFASRHGKRHGKSRKRLRCSKKPLHVNFKELGWDWIITAPLEYEAYHCEGVCDPPLRSH 394
DB 375 RAPLANRQGRKPSKALHVNFKDMGWDDWIITAPLEYEAFHCEGLCEPLRSH 434
QY 395 LEPTNHAIQTILMNSMDPGSTPPSCVPTKLTPIISILYIDAGNNVYKQYEDMVVYESGCCR 454
DB 435 LEPTNHAIQTILMNSMDPESTPTCCVPTRLSPISILFIDSANNVYKQYEDMVVYESGCCR 494
QY 455 R 455
DB 495 R 495

RESULT 4

A55452
cartilage-derived morphogenetic protein 1 precursor - human
C:Species: Homo sapiens (man)
C>Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 17-Mar-2000
C:Accession: A55452
R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko
J. Biol. Chem. 269, 28227-28234, 1994
A>Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr
A:Reference number: A55452; MUID:95050604; PMID:7961761
A:Accession: A55452
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-501 <CHA>
A:Cross-references: GB:U13660; NID:g600731; PID:g600732
C:Genetics:
A:Gene: GDB:CDMP1
A:Cross-references: GDB:438940
C:Superfamily: inhibin

Query Match 37.5%; Score 918.5; DB 2; Length 501;
Best Local Similarity 44.3%; Pred. No. 1.5e-57;
Matches 207; Conservative 64; Mismatches 101; Indels 95; Gaps 15;

QY 38 GSTKGMRSRKEGKQMRAPRDSAGREGQPPQPRQDEPRAQ----- 79
DB 114 RTVTPKGQLPGGKASKAGSAPSSFLKKTRFPGTPREPKEFPPTTPHEYMLSLYRT 173

C>Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 19-May-2000
C/Accession: S43296
R/Storm, E.E.; Huynh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.
Nature 368, 639-643, 1994

A>Title: Limb alterations in brachypodism mice due to mutations in a new member of the
A/Reference number: S43294; MUID:94195427; PMID:8145850
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-151 <STO>
A/Cross-references: GB:U08339; NID:g488465; PIDN:AAA18780.1; PID:g488466
C/Superfamily: inhibin

Query Match 22.4%; Score 549; DB 2; Length 151;
Best Local Similarity 59.5%; Pred. No. 5.9e-32;
Matches 103; Conservative 13; Mismatches 35; Indels 22; Gaps 1;

QY 283 RSQRKNLFAMREOLGSAEAGPCAGACGGSWPPSGAPDARPWLPSPGRRRRTAFASRH 342
DB 1 RRRRTTALCTRGAGSQQGGGGGGGGGGGGGA----- 38

QY 343 GKRGHKKSRKCKKLHVNFKELGWDWIIATPLEYAEYHCVGCDPLRLSHLEPTNHAI 402
DB 39 GRGHGRGRGRSCRKSUHVFKELGWDWIIATPDYBAYCEGVCDPFLRLSHLEPTNHAI 98

QY 403 IQTLNMNDPGSTPPSCVPTKTLPISILYIDANNVVYKYEDVMVYESGCGR 455
DB 99 IQTLNLSMAPDAAPASCVCVPARLSPISILIYIDANNVVYKYEDVMVVEACGR 151

RESULT 7
JH0687
bone morphogenetic protein 2I precursor - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C/Accession: JH0687; S16244
R/Nishimatsu, S.; Suzuki, A.; Shoda, A.; Murakami, K.; Ueno, N.
Biochem. Biophys. Res. Commun. 186, 1487-1495, 1992
A>Title: Genes for bone morphogenetic proteins are differentially transcribed in earl
A/Reference number: JH0687; MUID:92378616; PMID:1510675
A/Accession: JH0687
A/Molecule type: mRNA
A/Residues: 1-398 <NIS>
A/Cross-references: GB:X63424; NID:g64585; PIDN:CAA45018.1; PID:g64586
A/Experimental source: oocyte
R/Plessow, S.; Koester, M.; Knoechel, W.
Biochim. Biophys. Acta 1089, 280-282, 1991
A>Title: cDNA sequence of Xenopus laevis bone morphogenetic protein 2 (BMP-2).
A/Reference number: S16244; MUID:91274367; PMID:2054389
A/Accession: S16244
A/Molecule type: mRNA
A/Residues: 1-6,'S','8-15,'V',17-232,'N',234-398 <LEE>
A/Cross-references: EMBL:X55031; NID:g64581; PIDN:CAB38850.1; PID:g64582
C/Superfamily: inhibin
C/Keywords: dimer; glycoprotein
F/285-398/Product: bone morphogenetic protein 2I status predicted <MAT>
F/137,202,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.2%; Score 470; DB 2; Length 398;
Best Local Similarity 31.1%; Pred. No. 6.4e-26;
Matches 125; Conservative 60; Mismatches 117; Indels 100; Gaps 17;

QY 86 PGRGPRVWPHEYMILSIPTY--STAELKGINAFFQSXS---ANITSFVDGLDLHS 140
DB 65 PTPEGNNVIPPYMLDLHLHLAQADEGTSMDFQMERAASRANTVSFHFE--ESMEE 122

QY 141 TPLRR----OKYLFDVMSLDKDELVAERLERFOAPSAPGWPPAGPLH-VOLFPCLSPL 195
DB 123 IPSREKTIOGFNLSSINPELVTSALERIFREQVOEPFSDSKLHRINIYDIVKPA 182

QY 196 -----LLDARTLDPOGAPPAGVEFVDVWGQL-----RHQPWKQLCLELRANGEL 240
DB 183 AAASRGDPWRILTIRIVHHN---FSKWSEFSDIPDATAPMTAKDNHGCFVEVT---HT 235

```
QY 241 DAGEAARAGQPPPPDLRLSLGFGRRVRP-----PQERALLVVFTRSORKNLFAEMRE 295
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 236 D-----NDKNYPKKHVR---LSRLTPDKDNWQPIRLPLVTFESH----- 272
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 296 QLGSAFAAGAGAGSWPPSGADARWLPSPGRRRRRTAFASRGHKKRKGKSLRL-- 353
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 -----GKHALH-----KROKRA-----RHKORRLKSS 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 CSKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIOTLMSMDPG 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 CRHPLVYDFSDVGNWDWIVAPGGYHAFYCHGECFPPLADHLNSTNHAIOTLVNSVNT- 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 STPPSCCVPTKLTIPISILYIDAGNNVVKQYEDMVVESCGR 455
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 NIPKACCVPTLSAISMVLDENKVLKNYQDMVVEGCGCR 398

RESULT 8
JH0688
bone morphogenetic protein 2II precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: JH0688
R:Nishimatsu, S.; Suzuki, A.; Shoda, A.; Murakami, K.; Ueno, N.
Biochem. Biophys. Res. Commun. 186, 1487-1495, 1992
A:Title: Genes for bone morphogenetic proteins are differentially transcribed in early a
A:Reference number: JH0687; MUID:92378616; PMID:1510675
A:Accession: JH0688
A:Molecule type: mRNA
A:Residues: 1-398 <NIS>
A:Cross-references: GB:X63425; NID:g64583; PIDN:CAA45019.1; PID:g64584
A:Experimental source: oocyte
C:Superfamily: inhibin
C:Keywords: glycoprotein
F:285-399/Product: bone morphogenetic protein 2II #status predicted <MAP>
F:137,202,327,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

      Query Match      19.0%; Score 464; DB 2; Length 398;
      Best Local Similarity 30.1%; Pred. No. 1.7e-25;
      Matches 120; Conservative 68; Mismatches 117; Indels 94; Gaps 16;

QY 86 PGGRPRVVPHEYMYSIYRTYS--IAEKLG---INASFFQSSKSANTITSF-VDRGLDDLS 139
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 PTPGKNVIPPYMLDYLHLHSAQLADDOGSSEVDYHMERAAASRANTVRSFHEESMEETP 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 HTPLRR-QKYLFDVMSLSDKELVGAELFLFQAAPSAPWGPAGPLH-VQLEPPCLSP-- 195
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 125 ESGETIQRFNLSIPDEELVTSSELRIREFQVQEPFKTDGSKLHRIINYDIKPAAA 184
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 196 -----LLDARTLDPOGAPGAGVEFDVMQGL-----RHOPWKQLCLELRAANGELDA 242
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 ASRGVPVRLDRLTHHN---ESKVESFDVTPATIRWTAHKQPNHGFFVEVTHLNDNDNV 241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 GAEARAGQPPPPDLRLSL-----GFGRRVRPQERALLVVFTRSORKNLFAEMREQLG 298
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 242 PKRHVRIS-----RSLDLKGHWPRIRP-----LLVTFESH----- 272
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 299 SAEAGAGAGAGSWPPSGADARWLPSPGRRRRRTAFASRGHKKRKGKSLRL--CSK 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 -----GKHALH-----KROKRA-----RHKORRLKSSCR 300
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 KPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIOTLMSMDPGGTP 416
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 301 HPLVYDFSDVGNWDWIVAPGGYHAFYCHGECFPPLADHLNSTNHAIOTLVNSVNT-NIP 359
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 417 PSCCVPTKLTIPISILYIDAGNNVVKQYEDMVVESCGR 455
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 KACCVPTLSAISMVLDENKVLKNYQDMVVEGCGCR 398

RESULT 9
150608
bone morphogenetic protein 4 - chicken
C:Species: Gallus gallus (chicken)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: 150608
R:Francis, P.H.; Richardson, M.K.; Brickell, P.M.; Tickle, C.
Development 120, 209-218, 1994
A:Title: Bone morphogenetic proteins and a signalling pathway that controls patternin
A:Reference number: 150607; MUID:94163974; PMID:8119128
A:Accession: 150608
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-405 <FRA>
A:Cross-references: EMBL:X75915; NID:g472929; PIDN:CAA53514.1; PID:g472930
C:Genetics:
A:Gene: Bmp-4
C:Superfamily: inhibin

      Query Match      18.5%; Score 453; DB 2; Length 405;
      Best Local Similarity 28.8%; Pred. No. 1e-24;
      Matches 138; Conservative 69; Mismatches 162; Indels 110; Gaps 20;

QY 7 LLSAVFLISFLWDLPGFOQASISSCSAEGLSTKGM-----RSRKEGKMORAPRDSAG 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 MLMVILLCQVL--LGGTNHASLIPETGRKVAELQOAGSGRRSAQSHELLRGFFETLIQ 64
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 REGQEPQRPQDEPRAQOPRAQEPGRGPRVVPHEYMYSIYRTYSIAEK---LGINASF 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 MFLRRRPQPSKS-----AVIP-SYMLDLYRLQSGEEERSLQELISLQY 107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 118 FOSSKS-ANTITSF-VDRGLDDL---SHTPLRRQKYLFDVMSLSDKELVGAELRLFR-- 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 108 PERSASRANTVRSFHEESLVPGPSEAP--RIRFENLVSSVPDNEVISSELRLYREQ 165
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 -QAPSAPWGPAGPLHVLQPLCLSP-----LLDARTLDPOGAPGAGVEFDVMQG 220
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 VEEPSAAN--ERGFRIINIEYMKPLSRSAITRLDTRLVHHN---VTRMETFDV--- 217
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 221 LRHQPWKQLCLELRAANGELDAEAEARAGQPPPPDLRLSLGFGRRVRPQERALLV 280
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 -----SPAVIRWTKDKQP-----NHGLVIE 237
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 281 FTRSORKNLFAEMRDLQSGAEAGAGAGSWPPSGADARWLPSP--PGRRRRTA 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 VTHLHQATHQGHKRVIRSKSLPQGHG---GDW-----AQLRPLLVTFGHGDRGHALTR 287
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 338 FASRGCKRHG--KKSRLRCSKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLE 396
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 RARRSPKHGSRKKNKNCRRHAYVDFSDVGNWDWIVAPGGYAFYCHGDCFPPLADHLN 347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 397 PTNHAIOTLMSMDPGSTPPSCCVPTKLTIPISILYIDAGNNVVKQYEDMVVESCGR 455
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 STNHAIVOTLVNSVN-SSIPRACCVPTLSAISMVLDYDKVVLKNYQDMVVEGCGCR 405
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
BMH04
bone morphogenetic protein 4 precursor - human
N:Alternate names: bone morphogenetic protein 2B
C:Species: Homo sapiens (man)
C>Date: 16-Sep-1992 #sequence_revision 03-Aug-1995 #text_change 18-Jun-1999
C:Accession: C37278
R:Wozney, J.M.; Rosen, V.; Celeste, A.J.; Mitsock, L.M.; Whitters, M.J.; Kriz, R.W.;
Science 242, 1528-1534, 1988
A:Title: Novel regulators of bone formation: molecular clones and activities.
A:Reference number: A37278; MUID:89072730; PMID:3201241
A:Accession: C37278
A:Molecule type: mRNA
A:Residues: 1-408 <WO3>
A:Cross-references: GB:M22490; NID:g179503; PIDN:AAA51835.1; PID:g179504
C:Genetics:
A:Gene: GDB: BMP4; BMP2B
A:Cross-references: GDB:125205; OMIM:112262
A:Map position: 14q22-14q23
```


C:Superfamily: inhibin

C:Keywords: bone; glycoprotein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-292/Domain: propeptide #status predicted <PRO>

F:293-408/Product: bone morphogenetic protein 4 #status predicted <MAT>

F:143,208,350,365/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.0%; Score 441.5; DB 1; Length 408;

Best Local Similarity 28.0%; Pred. No. 6.8e-24;

Matches 135; Conservative 70; Mismatches 164; Indels 113; Gaps 19;

QY 7 LLSAVFLISFLWDLPGFQQAISISSSSCAELGSKG-MRSRKEGKMQRAPRDSDA---GR 62

DB 7 MLMVLLCOVL--LGGASHASLIPETGKKVAIOHAGGRSGQSHELLRDFEATLLQM 64

QY 63 EGQEPQRPQDEPRAQPPGPRVVPHEYMLSIVTYTYSIAEK-----LGINAS 116

DB 65 FGLRRRPQPSKS-----AVIP-DYMRDLYRLQSGEEEEEIHSTGLEVP 107

QY 117 FFOSSKANTITSF-VDRGLDLSHTPLRQ-KYLFVDSMLSDKEELVGAELRLFRQAPS 174

DB 108 ERPASR-ANTVRSFHHEHLENIPGTSSENSAFRLFNLSIPENEVISSAELRLFRE--Q 164

QY 175 APWGP-PAGPLHVQLFPCLSP-----LLDARTLDPOGAPPAGWEVDFVWQGL 221

DB 165 VQGGPWEGFHRINIEYVMKPPAEVVPGHILITRLDTRLVHHN--VTRWETFDV--- 217

QY 222 RHQPMKQLCLELRAANGELDAGEAARAGPQPPDLRLSLGFGRRVRPPQERALLVVF 281

DB 218 -----SPAVLR---WTREKQPNYGLATEVTH 240

QY 282 TRSQKNLFAEMREQLGSAEAGAGAGSGWPPPGAPDARPLWSPGR-----RR 333

DB 241 LHOTRTHOGHVIRISLSPOGS-----GNW-----AQLRPLLVTFGHDGRHALTR 287

QY 334 RRTAFASRHKRGKRSKRLCKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDPLRS 393

DB 288 RRAKSPKHSQARAKKKNCRHSLSYDFSDVGNWDWIVAPPGYQAFYCHGDCPPFLAD 347

QY 394 HLPPTNHAIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMVYESCG 453

DB 348 HLNSTNHAIQTLVNSVN-SSIPKACCVPTLSAISMLYLDYDKVVLKNYQEMVVEGCG 406

QY 454 CR 455

DB 407 CR 408

RESULT 11

I49541

bone morphogenetic protein 4 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999

C:Accession: I49541; S29523; B34201

R:Feng, J.Q.; Chen, D.; Cooney, A.J.; Tsai, M.; Harris, M.A.; Tsai, S.Y.; Feng, M.; Mund

J. Biol. Chem. 270, 28364-28373, 1995

A:Title: The mouse bone morphogenetic protein-4 (BMP-4) gene: Analysis of promoter utiliza

A:Reference number: I49541; MUID:96081880; PMID:7499338

A:Accession: I49541

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-420 <RES>

A:Cross-references: GB:I47480; NID:g994733; PIDN:AA037698.1; PID:g994734

R:Dickinson, M.E.; van der Meer-de Jong, R.; Hogan, B.L.M.

submitted to the EMBL Data Library, December 1990

A:Description: Nucleotide sequence of the mouse Bone Morphogenetic Protein-4 (BMP-4) cDN

A:Reference number: S29523

A:Accession: S29523

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 13-420 <DIC>

A:Cross-references: EMBL:X56848; NID:g50180; PIDN:CAA40179.1; PID:g50181

R:Dickinson, M.E.; Kobrin, M.S.; Silan, C.M.; Kingsley, D.M.; Justice, M.J.; Miller, D.A

peland, N.G.; Jenkins, N.A.

Genomics 6, 505-520, 1990

A:Title: Chromosomal localization of seven members of the murine TGF-beta superfamily

A:Reference number: A34201; MUID:90228966; PMID:1970330

A:Accession: B34201

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

C:Genetics: A:Residues: 253-420 <DI2>

A:Gene: BMP-4

A:Introns: 11/1; 137/1

C:Superfamily: inhibin

Query Match 17.8%; Score 436.5; DB 2; Length 420;

Best Local Similarity 28.5%; Pred. No. 1.6e-23;

Matches 139; Conservative 62; Mismatches 163; Indels 123; Gaps 19;

QY 7 LLSAVFLISFLWDLPGFQQAISISSSSCAELGSKG-MRSRKEGKMQRAPRDSDA---GR 62

DB 19 MLMVLLCOVL--LGGASHASLIPETGKKVAIOHAGGRSGQSHELLRDFEATLLQM 76

QY 63 EGQEPQRPQDEPRAQPPGPRVVPHEYMLSIVTYTYSIAEK-----LGINASF 117

DB 77 FGLRRRPQPSKS-----AVIP-DYMRDLYRLQSGEEEEEQSGTGLEY 119

QY 118 FFOSSK-ANTITSFVDRGLDLSHTPLRQ---KYLFVDSMLSDKEELVGAELRLFR-- 170

DB 120 PERPASRANTVRSFHE--EHLNIPGTSSENSAFRLFNLSIPENEVISSAELRLFREQ 177

QY 171 --QAPSAPWG-----PPAG--PLHVQLFPCLSPDLLDARTLDPOGAPPAGWEVF 215

DB 178 VQGGPWEGFHRINIEYVMKPPAEVVPGH-----LITRLDTRLVHHN--VTRWETF 228

QY 216 DWQGLRHQHPWKQLCLELRAANGELDAGEAARAGPQPPDLRLSLGFGRRVRPPQER 275

DB 229 DV-----SPAVLR---WTREKQPNYGL 247

QY 276 ALIVVTRSORKNLFAEMREQLGSAEAGAGSGWPPPGAPDARPLWSPGR--- 331

DB 248 AIEVTHLHOTRTHOGHVIRISLSPOGSQDWA-----QLRPLLVTFGHDGRG 294

QY 332 ---RRRTAFASRHKRGKRSKRLCKKPLHVNFKELGWDWIIAPLEYAYHCEGVCD 388

DB 295 HTLTRRRKRSPKHPQSRKKNCRHSLSYDFSDVGNWDWIVAPPGYQAFYCHGDCP 354

QY 389 FPLRSHLEPTNHAIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMV 448

DB 355 FPLADHLNSTNHAIQTLVNSVN-SSIPKACCVPTLSAISMLYLDYDKVVLKNYQEMV 413

QY 449 VESCGCR 455

DB 414 VEGCGCR 420

RESULT 12

S37073

bone morphogenetic protein 2 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999

R:Feng, J.Q.; Chen, D.; Feng, M.; Harris, M.A.; Mundy, G.R.; Harris, S.E.

submitted to the EMBL Data Library, September 1993

A:Description: cDNA sequence of fetal rat calvarial osteoblast bone morphogenetic pro

A:Reference number: S37073

A:Accession: S37073

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-393 <FEN>

A:Cross-references: EMBL:225868; NID:g397950; PIDN:CAA81088.1; PID:g397951

C:Superfamily: inhibin

Query Match 17.7%; Score 434; DB 2; Length 393;

Best Local Similarity 30.2%; Pred. No. 2.2e-23;


```
QY 202 LDPQAGPAGWEVDFWQGLRHQPWKQLCLELRAANGELDAGAEARARGPQPPPPDLR 261
Db 197 VN-QNA--SRWESFDYTPAMVR--W-----TAQGHANHGFTVVEVAHLEEKQ----- 237
QY 262 SLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAFAAGPGAGAGSMPPPSGAPD 321
Db 238 --GVSKR-----HVRISRLHQD-----EHSW-----SQ 259
QY 322 ARPWLSPG-----RRRRTAFASRHGRHKGKSRLL--CSKKPLHVNFKELGWDD 370
Db 260 IRPLLVTFGHDGKGPLHKKRQA-----KHKQRKRLKSSCKRHPPLYVDVDFSDVGWND 312
QY 371 WIITAPLEYAYHCEGVCDFPLSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLPTPI 430
Db 313 WIVAPPGYHAFYCHGECPPFLADHLNSTNHAIIVQTLVNSVN-SKIPKACCVPTLSAISM 371
QY 431 LYIDAGNNVVYKQYEDMVVESCGR 455
Db 372 LYIDENEKVVLKYNQDMVVEGCGCR 396

RESULT 15
S52408
SPDVR1 protein - sea urchin (Strongylocentrotus purpuratus)
C:Species: Strongylocentrotus purpuratus (purple urchin)
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 20-Sep-1999
C:Accession: S52408
R:Ponce, M.R.; Micol, J.L.; Davidson, E.H.
submitted to the EMBL Data Library, February 1995
A:Description: SPDVR1, a member of the transforming growth factor-beta superfamily expe
A:Reference number: S52408
A:Accession: S52408
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <PON>
A:Cross-references: EMBL:248313; NID:g673496; PID:g673497
C:Superfamily: inhibin

Query Match 17.6%; Score 429.5; DB 2; Length 461;
Best Local Similarity 27.7%; Pred. No. 5.5e-23;
Matches 140; Conservative 68; Mismatches 177; Indels 121; Gaps 19;

QY 6 VLLSAVFLISFLWDL---PGFQOASISSCSAEALGSGTKGMRKRGKMQRAPRDSAGR 62
Db 19 LILSLFFGPGLSWDFYSGDEQDLSLARERRAANYNPSPHMTWERNETIQ-----EILNI 74
QY 63 EGQEPQPRP-----ODEPRAQ-----QPRAQEPGPGRPVVPHEYMLSIYRTYSTA 108
Db 75 LGLQHRPRPPLSRGGNQFCQFTWSYRTLNIDEQSGHPSETEPQPGGLASNAIYNSP 134
QY 109 EKLGINA-----SFFQSSKSANTITS---FVDRGLDLSHTPLRRQKVLFDVS 153
Db 135 DSSGIGSVMSGTVFNVTREVAQVQADTMSLPVHYKDAAIETDEH-----RYRFDIG 188
QY 154 MLSDEELVGAELRFRQAPSAPWPPAGPLHVQLF-----PCLSPLLLDARTLDPOGA 207
Db 189 RIPOGETVTSAEALRVFERDA--GROGRSLYRIDVLLRERSGDSRSPVYLDSTIV--CA 243
QY 208 PPAGWEVDV-----WQGLRHQPWKQLCLELRAANGELDAGAEARARGPQPPPPDLR 261
Db 244 GDHGLVFDMTSATSTW---RSYPGANVGLQURVE--SLQGLNID-----PTDAG 288
QY 262 SLGFGRRVRPPQERALLVVFTRSQ-----RKNLFAEMREQLGSAFAAGPGAGAGS 312
Db 289 VYGVGN--NEGPEPWVFFQNEEVIAATNSHLRNRRAATQKKG----- 332
QY 313 WPPPSGAPDARWLPSPGRRRR---TAFASRHGRHKGKSRLLRCSKKPLHVNFKELGW 368
Db 333 -----GKRPRKPTDNDIASR-DSASSLNSDWQCKRKNLFVNFEEDLDW 374
QY 369 DDWIIITAPLEYAYHCEGVCDFPLSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLPTI 428
Db 375 QEWIIITAPLEYAYHCEGVCDFPLSHLEPTNHAIQTLMNSMDPGSTPPSCCVTKLPTI 434
```

```
QY 429 SILYIDAGNNVVYKQYEDMVVESCGR 454
Db 435 TVLYYDDSRNVVLKYYKKNMVVRACGC 460
```

Search completed: November 25, 2002, 02:58:59
Job time : 52 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:54:00 ; Search time 36 seconds
(without alignments)
371.873 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLWDL.....GNVVYKQVDMVVEGCGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2.6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2.6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2.6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2.6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2.6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2.6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1747	71.4	321	1	US-08-362-670B-26
2	1747	71.4	321	3	US-08-333-576C-26
3	1747	71.4	321	4	US-08-808-324-26
4	1747	71.4	321	5	PCT-US94-14030A-26
5	1202.5	49.1	263	1	US-08-362-670B-32
6	1202.5	49.1	263	3	US-08-333-576C-32
7	1202.5	49.1	263	4	US-08-808-324-32
8	1202.5	49.1	263	5	PCT-US94-14030A-32
9	927.5	37.9	501	2	US-08-288-508C-2
10	919.5	37.6	495	1	US-08-455-559-10
11	919.5	37.6	495	4	US-09-145-060-10
12	919.5	37.6	495	5	PCT-US94-00657-10
13	917.5	37.5	401	3	US-08-289-222E-3
14	917.5	37.5	401	4	US-09-054-528B-3
15	765	31.3	388	1	US-08-362-670B-34
16	765	31.3	388	3	US-08-333-576C-34
17	765	31.3	388	4	US-08-808-324-34
18	765	31.3	388	5	PCT-US94-14030A-34
19	731	29.9	134	1	US-08-581-529B-6
20	731	29.9	134	3	US-09-097-616-6
21	731	29.9	134	5	PCT-US94-07762-6
22	691.5	28.3	294	1	US-08-362-670B-2
23	691.5	28.3	294	3	US-08-333-576C-2
24	691.5	28.3	294	4	US-08-808-324-2
25	691.5	28.3	294	5	PCT-US94-14030A-2
26	661	27.0	119	1	US-08-581-529B-7
27	661	27.0	119	3	US-09-097-616-7

28	661	27.0	119	5	PCT-US94-07762-7	Sequence 7, Appl
29	627	25.6	411	1	US-08-362-670B-28	Sequence 28, Appl
30	627	25.6	411	3	US-08-333-576C-28	Sequence 28, Appl
31	627	25.6	411	4	US-08-808-324-28	Sequence 28, Appl
32	627	25.6	411	5	PCT-US94-14030A-28	Sequence 28, Appl
33	618	25.3	240	1	US-08-362-670B-30	Sequence 30, Appl
34	618	25.3	240	3	US-08-333-576C-30	Sequence 30, Appl
35	618	25.3	240	4	US-08-808-324-30	Sequence 30, Appl
36	618	25.3	240	5	PCT-US94-14030A-30	Sequence 30, Appl
37	553	22.6	161	2	US-08-581-528A-6	Sequence 6, Appl
38	553	22.6	161	5	PCT-US94-07799-6	Sequence 6, Appl
39	549	22.4	119	1	US-08-455-559-13	Sequence 13, Appl
40	549	22.4	119	4	US-09-145-060-13	Sequence 13, Appl
41	549	22.4	119	5	PCT-US94-00657-13	Sequence 13, Appl
42	549	22.4	120	1	US-08-362-670B-4	Sequence 4, Appl
43	549	22.4	120	3	US-08-333-576C-4	Sequence 4, Appl
44	549	22.4	120	4	US-08-808-324-4	Sequence 4, Appl
45	549	22.4	120	5	PCT-US94-14030A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-362-670B-26
; Sequence 26, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-670B-26

Query Match 71.4%; Score 1747; DB 1; Length 321;

Best Local Similarity 100.0%; Pred. No. 4.3e-141;

Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRROKYLFDYMSLSKDEELVGAELRFQAPSPAGPLHVQLFPCLSPLL 196

|||||

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSAPWGPAGPLHVQLFPCLSPILL 62
QY 197 LDARTLDPOGAPAGWEVDVWQGLRHQPKQLCLELRAAWGELDAGEAEARAGPQQPP 256
Db 63 LDARTLDPOGAPAGWEVDVWQGLRHQPKQLCLELRAAWGELDAGEAEARAGPQQPP 122
QY 257 PDLRLSLGFRGRRVVRPQERALLVVFTRRSQKKNLFAEMRQOLGSAEAGPGAGAGSWPPP 316
Db 123 PDLRLSLGFRGRRVVRPQERALLVVFTRRSQKKNLFAEMRQOLGSAEAGPGAGAGSWPPP 182
QY 317 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKPLHVNFKELGWDDWIIAPL 376
Db 183 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKPLHVNFKELGWDDWIIAPL 242
QY 377 EYEAVHCEGVCDFPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPTISILYIDAG 436
Db 243 EYEAVHCEGVCDFPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPTISILYIDAG 302
QY 437 NNVYKQYEDMVESGCR 455
Db 303 NNVYKQYEDMVESGCR 321

RESULT 2

US-08-333-576C-26
; Sequence 26, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333.576C
; FILING DATE: No. 6027919ember 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-576C-26

Query Match 71.4%; Score 1747; DB 3; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSAPWGPAGPLHVQLFPCLSPILL 196
|||||

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSAPWGPAGPLHVQLFPCLSPILL 62
QY 197 LDARTLDPOGAPAGWEVDVWQGLRHQPKQLCLELRAAWGELDAGEAEARAGPQQPP 256
Db 63 LDARTLDPOGAPAGWEVDVWQGLRHQPKQLCLELRAAWGELDAGEAEARAGPQQPP 122
QY 257 PDLRLSLGFRGRRVVRPQERALLVVFTRRSQKKNLFAEMRQOLGSAEAGPGAGAGSWPPP 316
Db 123 PDLRLSLGFRGRRVVRPQERALLVVFTRRSQKKNLFAEMRQOLGSAEAGPGAGAGSWPPP 182
QY 317 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKPLHVNFKELGWDDWIIAPL 376
Db 183 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKPLHVNFKELGWDDWIIAPL 242
QY 377 EYEAVHCEGVCDFPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPTISILYIDAG 436
Db 243 EYEAVHCEGVCDFPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPTISILYIDAG 302
QY 437 NNVYKQYEDMVESGCR 455
Db 303 NNVYKQYEDMVESGCR 321

RESULT 3

US-08-808-324-26
; Sequence 26, Application US/08808324
; Patent No. 6284872
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808.324
; FILING DATE: Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-808-324-26

Query Match 71.4%; Score 1747; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSAPWGPAGPLHVQLFPCLSPILL 196
|||||

Db 3 DLSTPLRRQKYLFDVSMYSDKELVGAELRLFRQAPSWGPPAGPLHVQLPCLSPLL 62
QY 197 LDARTLDPOGAPGAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOQPP 256
Db 63 LDARTLDPOGAPGAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOQPP 122
QY 257 PDLRSGLGFGRRVPPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPGAGAGSWPPP 316
Db 123 PDLRSGLGFGRRVPPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPGAGAGSWPPP 182
QY 317 SGAPDARPLSPGRRRRRTAFASRGKRGKSKKPLHVNFKELGWDWIIAPL 376
Db 183 SGAPDARPLSPGRRRRRTAFASRGKRGKSKKPLHVNFKELGWDWIIAPL 242
QY 377 EYAYHCEGVCDPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAG 436
Db 243 EYAYHCEGVCDPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAG 302
QY 437 NNVYKQYEDMVVESCGR 455
Db 303 NNVYKQYEDMVVESCGR 321

RESULT 4

PCT-US94-14030A-26
; Sequence 26, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-14030A-26

Query Match 71.4%; Score 1747; DB 5; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSTPLRRQKYLFDVSMYSDKELVGAELRLFRQAPSWGPPAGPLHVQLPCLSPLL 196
Db 3 DLSTPLRRQKYLFDVSMYSDKELVGAELRLFRQAPSWGPPAGPLHVQLPCLSPLL 62
QY 197 LDARTLDPOGAPGAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOQPP 256
Db 63 LDARTLDPOGAPGAGVEFDVWQGLRHQPKQLCLELRAAWGELDAGEAARARGPOQPP 122
QY 257 PDLRSGLGFGRRVPPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPGAGAGSWPPP 316
Db 123 PDLRSGLGFGRRVPPQERALLVVFTRSQKKNLFAEMREQLGSAEAGPGAGAGSWPPP 182
QY 317 SGAPDARPLSPGRRRRRTAFASRGKRGKSKKPLHVNFKELGWDWIIAPL 376
Db 183 SGAPDARPLSPGRRRRRTAFASRGKRGKSKKPLHVNFKELGWDWIIAPL 242
QY 377 EYAYHCEGVCDPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAG 436
Db 243 EYAYHCEGVCDPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAG 302
QY 437 NNVYKQYEDMVVESCGR 455
Db 303 NNVYKQYEDMVVESCGR 321

RESULT 5

US-08-362-670B-32
; Sequence 32, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-670B-32

Query Match 49.1%; Score 1202.5; DB 1; Length 263;
Best Local Similarity 83.3%; Pred. No. 7.7e-95;
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;

Qy	240	LDAGEAEARAGQPQQPPDLSLGFGRVRRVPOERALLVVFTRSQKNLFAEMREQLGS	2999
Db	52	LDXGDTGARARGQPQPPDLSLGFGRVRRVPOERALLVVFTRSQKNLFTEMHEQLGS	1111
Qy	300	AEAAGPCAGAEGSWPPSGAPDARPMLSPGRRRRRTAFASRGRGKRGKKSRLRCSKKPL	3599
Db	112	AEA-----AGAEGSCAPSGSPDTSGLPSPGRRRRRTAFASRGRGKRGKKSRLRCSKKPL	1677
Qy	360	HYNFKELGWDWIIAPLEYEAYHCEGVCDDPLRSHLEPTNHAIQTLMNSMDPGSTPPSC	4199
Db	168	HYNFKELGWDWIIAPLEYEAYHCEGVCDDPLRSHLEPTNHAIQTLMNSMDPGSTPPSC	2277
Qy	420	CVPTKLPPIISILYIDAGNNVVYKQYEDMVYESGCR	455
Db	228	CVPTKLPPIISILYIDAGNNVVYKQYEDMVYESGCR	263

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 49.1%; Score 1202.5; DB 4; Length 263;
Best Local Similarity 83.3%; Pred. NO. 7.7e-95;
Matches 230: Conservative 6; Mismatches 21; Indels 19; Gaps 4.

QY	180	PAGLHVLQFPCULSPDLLDARTLDPOGAPPAGWEVDFVWQGLRHQPWKQJCLLELRAAWGE	239
Db	7	PAG-----PTLRSSGTPQR---PAG-KSEDFVWQGLRHQPWKQJCLLELRAAWGE	51
QY	240	LDAGEAEARAGPQQQPPPPDLRLSLGFGFRRVRPQERALLVFTFSORKNLFLEMRRLQSGS	299
Db	52	LDXGDTGARARGPQQPPDLRLSLGFGFRRVRPQERALLVFTFSORKNLFLEMRRLQSGS	111

QY 300 AEAAGCAGAGSWPPSPGAPDARPNLPSGRRRRRTAFASRHGKRGKSRKPL 359
Db 112 AEA-----AGAEGSCAPSGSPDGTGSLWLPSPGRRRRRTAFASRHGKRGKSRKPL 167
QY 360 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSC 419
Db 168 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSC 227
QY 420 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 455
Db 228 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 263

RESULT 8

PCT-US94-14030A-32
; Sequence 32, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-14030A-32

Query Match 49.1%; Score 1202.5; DB 5; Length 263;
Best Local Similarity 83.3%; Pred. No. 7.7e-95;
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;
QY 180 PAGPLHVQLPCLSPDLLDARTLDPOGAPGAGVEFDVMOGLRHQPWKQLCLELRAWGE 239
Db 7 PAG-----PVLRGSSGTQPR---PAG--KSFVMOGLRPQWKQLCLELRAWGE 51
QY 240 LDAGEAEARAGQQPPPDLSLGFGRVRRPQERALLVVFTRSKNLFASMRQGLS 299
Db 52 LDAGDTGARAGQQPPPDLSLGFGRVRRPQERALLVVFTRSKNLFTEMHQGLS 111
QY 300 AEAAGCAGAGSWPPSPGAPDARPNLPSGRRRRRTAFASRHGKRGKSRKPL 359
Db 112 AEA-----AGAEGSCAPSGSPDGTGSLWLPSPGRRRRRTAFASRHGKRGKSRKPL 167

Db 112 AEA-----AGAEGSCAPSGSPDGTGSLWLPSPGRRRRRTAFASRHGKRGKSRKPL 167
QY 360 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSC 419
Db 168 HVNFKELGWDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSC 227
QY 420 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 455
Db 228 CVPTKLTPIISILYIDAGNNVVYKQYEDMVVESCGR 263

RESULT 9

US-08-288-508C-2
; Sequence 2, Application US/08288508C
; Patent No. 5994094
; GENERAL INFORMATION:
; APPLICANT: H tten, Gertrud
; APPLICANT: Neidhardt, Helge
; APPLICANT: Paulista, Michael
; TITLE OF INVENTION: NEW GROWTH/DIFFERENTIATING FACTOR OF
; TITLE OF INVENTION: THE TGF- FAMILY
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,508C
; FILING DATE: 10-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 43 26 829.3
; FILING DATE: 10-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 18 222.8
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 20 157.5
; FILING DATE: 09-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: JAHNS, Kristina M.
; REGISTRATION NUMBER: P-41,092
; REFERENCE/DOCKET NUMBER: P564-4019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-288-508C-2

Query Match 37.9%; Score 927.5; DB 2; Length 501;
Best Local Similarity 44.7%; Pred. No. 4.6e-71;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;
QY 38 GSTKGRMSRKGRKMQRPDSDAGREGQEPQPRPQDEPRAQQ----- 79
Db 81 GQTGGLTQPKDKPKLP-----PRPG-GPEPKGHPPTQRTQATARTVTPKGOLPGKAP 134
QY 80 -----PRAQEPGGRGPRVVPHEYMLSIYRTYSIAEKLGINASFFQ 119
Db 135 PKAGSVPSFLLKKAREPGPPPEPKPEFRPPPTTPHEYMLSLYRTLSDADKGGNSVKL 194

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-145-060-10

Query Match 37.6%; Score 919.5; DB 4; Length 495;
Best Local Similarity 43.5%; Pred. No. 2.2e-70;
Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PGFOASISSCSAELGSKGMRKQAPRDSADAGREGQPPQPPQDE-----GSETKPGSSQTRQAAA 113
DB 60 PGHHYGVGATNARAK-GSSGQTQAKKDEPKMPPRSG-----GSETKPGSSQTRQAAA 113
QY 75 -----PRAQOP-----RAQEP-----PGRGPRVVPHEYMLSYRT 104
DB 114 RTVTPKQGLPGGKASSKAGSAPSSFLKKTREPGTPREPKEPPRPPTITPHEYMLSYRT 173
QY 105 YSTAELGINASFQSSKSNANTITSFVDRGLDLSHTPLRRQKYLFDVMSLSDKEELVGA 164
DB 174 LSDADRGKGNSSVKLEAGLANTITSFIDKQDDRG-PAVRKQRYVFDISAL-EKDGLLGA 231
QY 165 ELRLFRQAP---SAPWGPAGPL-HVOLFPCLs-----PLLLDARTLDPQAGPAGWEVFD 216
DB 232 ELRLRRKPLDVAKPAVPSSGRVAQLKSSCPGQPAALLDVRV--PGLDGSQWEVFD 289
QY 217 VWOGLRH-QPWKQLCLELRAWGLDAGEAEARAGQPPPPDLRLSLGRRVRRPQER 275
DB 290 IWLFRNFKNSAQLCLELE-AW-----ERGRA-----VDRLGLGFERTARQVHEK 333
QY 276 ALLVVFTRSRKNI-FAEMREQLGSAAGPGAGAGSWPPSPGAPDAPRWLSPGRRRR 334
DB 334 ALFLVFGRTKKRDLFFNEIKARSGQDDKT-----VYELFSQ-RRKR 374
QY 335 RTAFASGRHGRKSKRLRCRKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFFPLRSH 394
DB 375 RAPLANRQGRKPSKRLKARCSRKALHVNFKDMGWDDWIIAPLEYAEPHCEGLCEFFPLRSH 434
QY 395 LEPTNHAIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMVVESC GC 454
DB 435 LEPTNHAVIQTLMNSMDPESTPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESC GC 494
QY 455 R 455
DB 495 R 495

RESULT 12
PCT-US94-00657-10
; Sequence 10, Application PC/TUS9400657
; GENERAL INFORMATION:
; APPLICANT: SE-JIN LEE
; APPLICANT: HUYNH, THANH
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1800 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00657
; FILING DATE: 1/12/94

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: WETHERELL, JR. PH.D., JOHN R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: FD3256 CIP OF PD2280
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-00657-10

Query Match 37.6%; Score 919.5; DB 5; Length 495;
Best Local Similarity 43.5%; Pred. No. 2.2e-70;
Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PGFOASISSCSAELGSKGMRKQAPRDSADAGREGQPPQPPQDE-----GSETKPGSSQTRQAAA 113
DB 60 PGHHYGVGATNARAK-GSSGQTQAKKDEPKMPPRSG-----GSETKPGSSQTRQAAA 113
QY 75 -----PRAQOP-----RAQEP-----PGRGPRVVPHEYMLSYRT 104
DB 114 RTVTPKQGLPGGKASSKAGSAPSSFLKKTREPGTPREPKEPPRPPTITPHEYMLSYRT 173
QY 105 YSTAELGINASFQSSKSNANTITSFVDRGLDLSHTPLRRQKYLFDVMSLSDKEELVGA 164
DB 174 LSDADRGKGNSSVKLEAGLANTITSFIDKQDDRG-PAVRKQRYVFDISAL-EKDGLLGA 231
QY 165 ELRLFRQAP---SAPWGPAGPL-HVOLFPCLs-----PLLLDARTLDPQAGPAGWEVFD 216
DB 232 ELRLRRKPLDVAKPAVPSSGRVAQLKSSCPGQPAALLDVRV--PGLDGSQWEVFD 289
QY 217 VWOGLRH-QPWKQLCLELRAWGLDAGEAEARAGQPPPPDLRLSLGRRVRRPQER 275
DB 290 IWLFRNFKNSAQLCLELE-AW-----ERGRA-----VDRLGLGFERTARQVHEK 333
QY 276 ALLVVFTRSRKNI-FAEMREQLGSAAGPGAGAGSWPPSPGAPDAPRWLSPGRRRR 334
DB 334 ALFLVFGRTKKRDLFFNEIKARSGQDDKT-----VYELFSQ-RRKR 374
QY 335 RTAFASGRHGRKSKRLRCRKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFFPLRSH 394
DB 375 RAPLANRQGRKPSKRLKARCSRKALHVNFKDMGWDDWIIAPLEYAEPHCEGLCEFFPLRSH 434
QY 395 LEPTNHAIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMVVESC GC 454
DB 435 LEPTNHAVIQTLMNSMDPESTPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESC GC 494
QY 455 R 455
DB 495 R 495

RESULT 13
US-08-289-222E-3
; Sequence 3, Application US/08289222E
; Patent No. 6120760
; GENERAL INFORMATION:
; APPLICANT: HOTTEN, GERTRUD
; APPLICANT: NEIDHARDT, HELGE
; APPLICANT: BECHTOLD, ROLF
; APPLICANT: POHL, JENS
; TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS OF THE TGF-B
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM
; STREET: 655 FIFTEENTH STREET, N. W., G STREET LOBBY,
; STREET: SUITE 330

CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: POHL, JENS
FILING DATE: 25-AUG-1999
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/289,222
FILING DATE: 12-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE P 44 23 190.3
FILING DATE: 07-JUL-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EPO 92102324.8
FILING DATE: 12-FEB-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00350
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KITTS, MONICA CHIN
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P564-9021
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 401 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-289-222E-3

Query Match 37.5%; Score 917.5; DB 3; Length 401;
Best Local Similarity 46.0%; Pred. No. 2.4e-70;
Matches 202; Conservative 59; Mismatches 87; Indels 91; Gaps 14;

QY 67 PQRPODEPRAQ-----PRAQPPGR 88
Db 4 PEPKPGHPQTRQATARTVTPKQPGKAPKAGSVPSFLLKKAREPGPPREKPEPR 63
QY 89 GRVVPHEYMLSYRTYSAEKLGINASFQSSKANTITSFVDRGLDLSHTPLRQKY 148
Db 64 PPTITPHEYMLSYRTLSADRGKGGNSVKLEAGLANTITSFIDKQDGRGPV-VRKQY 122
QY 149 LFDVMSLSKEELVGAELRFRQAPS-----APWGGPAGPLHVLPCLSLD-----PLLLD 198
Db 123 VEDISAL-ENKDLGLGAEELRLTRKPKSDTAKPAAPGGGAAQL--KLSSCPGSRQPSALLD 179
QY 199 ARTLDPQAGPAGVEFDVWQGLRH-OPWKQLCLELRAANGELDAGEAARAGPQPP 257
Db 180 VRSV--PGLDGSGWEFDLWKLFRNFKNSAQLCLELE-AW---ERGRA-----221
QY 258 PDLRSLGFRVRPPQERALLVFTTSQRKNL-FAEMRQGLGSAEAGPGAGAGSWPPP 316
Db 222 VDLRGLGFDRAARQVHEKALFLVFGRTKKRDLFFNEIKARSGDDKT-----268
QY 317 SGAPDARPLPSCRRRRRTAFASRKHGKHSRKLRLKARCSKPLHVNFKELGWDWIIAPL 376
Db 269 -----VYEFLSFO-RRKRRLATROGKRPSKNLKRCSKRLHVNFKDMGWDWIIAPL 322
QY 377 EYAFHCEGVCDFPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCVPTKLTPTISILYIDAG 436
Db 323 EYAFHCEGLCEPFLRSHLEPTNHAIIQTLNMSMDPESTPPTCVPTKLTPTISILYIDAG 382

QY 437 NNVMYKQYEDMVMVSCGR 455
Db 383 NNVMYKQYEDMVMVSCGR 401
RESULT 14
US-09-054-526B-3
Sequence 3, Application US/09054526B
Patent No. 6197550
GENERAL INFORMATION:
APPLICANT: H TTEN, GERTRUD
APPLICANT: NEIDHARDT, HELGE
APPLICANT: BECHTOLD, ROLF
APPLICANT: POHL, JENS
TITLE OF INVENTION: DNA SEQUENCES ENCODING NOVEL
TITLE OF INVENTION: GROWTH/DIFFERENTIATION FACTORS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM LLP
STREET: 655 FIFTEENTH STREET, N. W., G STREET LOBBY,
STREET: SUITE 330
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,526B
FILING DATE: 03-APR-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/289,222
FILING DATE: 12-AUG-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: DE P 44 23 190.3
FILING DATE: 01-JUL-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EPO 92102324.8
FILING DATE: 12-FEB-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/00350
FILING DATE: 12-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: KITTS, MONICA CHIN
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P564-8005
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 401 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-054-526B-3

Query Match 37.5%; Score 917.5; DB 4; Length 401;
Best Local Similarity 46.0%; Pred. No. 2.4e-70;
Matches 202; Conservative 59; Mismatches 87; Indels 91; Gaps 14;

QY 67 PQRPODEPRAQ-----PRAQPPGR 88
Db 4 PEPKPGHPQTRQATARTVTPKQPGKAPKAGSVPSFLLKKAREPGPPREKPEPR 63
QY 89 GRVVPHEYMLSYRTYSAEKLGINASFQSSKANTITSFVDRGLDLSHTPLRQKY 148
Db 64 PPTITPHEYMLSYRTLSADRGKGGNSVKLEAGLANTITSFIDKQDGRGPV-VRKQY 122

QY 149 LFDVSMLSKKEELVGAELRLFRAPS-----APWGPPAGPLHVQLPCLLS-----PILLD 198
Db 123 VFDISAL-EKQGLGAEELRLKRRKPSDTAKPAAPGGGAAQL--KLSCSPSGRQPASLLD 179
QY 199 ARTLDPQAGPAGWEVDFVWQGLRH-QPWKQLCLELRAAGELDAGEAEARARGPQQPPP 257
Db 180 VRSV--PGLDGSWEVDFVWQGLRH-QPWKQLCLELRAAGELDAGEAEARARGPQQPPP 221
QY 258 PDLRSIGRRVRPPQERALLVFTSRQKNL-FAEMREQLGSAEAGPAGAGSWPPP 316
Db 222 VDLRGLGDFRAARQVHEKALFLVFGRTKKRDLFFNEIKARSGDDKT----- 268
QY 317 SGAPDARPLPSGRRRRRTAFASRGKRGHKKSRKRLVHNFVKELGWDWIIAPL 376
Db 269 -----VYELFSQ-RKRRAPLATROCKRPSNLKARCSCRKALHVNFKDMGWDWIIAPL 322
QY 377 EYEAHCEGVCDFPLRSHLEPTNHAIQIOLMSMDPGSTCCVPTKLTPTISILYIDAG 436
Db 323 EYEAHCEGVCDFPLRSHLEPTNHAIQIOLMSMDPGSTCCVPTKLTPTISILYIDAG 382
QY 437 NNVYKQYEDMVVESCGR 455
Db 383 NNVYKQYEDMVVESCGR 401

RESULT 15
US-08-362-670B-34
; Sequence 34, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 388 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-670B-34

Query Match 31.3%; Score 765; DB 1; Length 388;
Best Local Similarity 51.2%; Pred. No. 2.3e-57;
Matches 172; Conservative 36; Mismatches 78; Indels 50; Gaps 10;

QY 146 QKYLFDVSMLSKKEELVGAELRLFRFRO-APSAPWGPAGPLHVQLPFC-----LSPLLLDAR 200
Db 77 QSFLEFDVSSINDADEVVGAEELRVLRGSPESGPGSWTSPPLLLSTCPGAARAPRLIYSR 136
QY 201 TLDPOGAPAGWEVDFVWQGLRH-----QPWKQLCLELRAAGELDAGEAEARARGPQQP 255
Db 137 AAEP--LVGQRWEAFVADAMRRHRREPFRPFACLLLRVAVGPV-----P 180
QY 256 PPDRLSLGF--GRRVRPPQERALLVFTSRQK-NLFAEMREQ---LGSAAEAGPAG 308
Db 181 SPLALRLGFGWPGGGGSAEAEAVLVVSRQKESLEFREIRAQARALGAALASEP--- 237
QY 309 AEGSWPPPAGDARWLPSPGRRRRRTAFASRH-----GKRHGKSRKRLCSKKPL 359
Db 238 -----LPDPCTGTASPRAVIGGRRRRRTALAGTRTAQGGGGAGRGHRRGRSRSRKPL 292
QY 360 HVNFVKELGWDWIIAPLEYEAHCEGVCDFPLRSHLEPTNHAIQIOLMSMDPGSTPPSC 419
Db 293 HVDFKELGWDWIIAPLEYEAHCEGVCDFPLRSHLEPTNHAIQIOLMSMDPGSTPPSC 352
QY 420 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 455
Db 353 CVPARLSPISILYIDAGNNVYKQYEDMVVESCGR 388

Search completed: November 25, 2002, 02:59:49
Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:55:26 ; Search time 30 Seconds
(without alignments)
237.533 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100480 seqs, 15661496 residues

Total number of hits satisfying chosen parameters: 100480

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1798	73.5	436	10	US-09-730-772-14
2	1798	73.5	436	10	US-09-735-849-14
3	1747	71.4	321	9	US-09-945-182-26
4	1202.5	49.1	263	9	US-09-945-182-32
5	927.5	37.9	501	8	US-08-981-490B-1
6	919.5	37.6	495	9	US-09-880-708-10
7	918.5	37.5	502	9	US-09-813-398-37
8	903.5	36.9	501	10	US-09-730-772-13
9	903.5	36.9	501	10	US-09-735-849-13
10	765	31.3	388	9	US-09-945-182-34
11	691.5	28.3	294	9	US-09-945-182-2
12	627	25.6	411	9	US-09-945-182-28
13	618	25.3	240	9	US-09-945-182-30
14	549	22.4	119	8	US-08-945-459A-1
15	549	22.4	119	9	US-09-880-708-13
16	549	22.4	119	9	US-09-068-253-2
17	549	22.4	120	9	US-09-945-182-4
18	442.5	18.1	408	10	US-09-749-728B-69
19	442.5	18.1	409	9	US-09-813-398-27

20	431.5	17.6	396	9	US-10-044-716-2
21	431.5	17.6	396	10	US-09-952-360-2
22	431.5	17.6	397	9	US-09-813-398-24
23	427	17.4	437	10	US-09-784-911-4
24	426	17.4	419	10	US-09-784-911-2
25	422	17.2	427	10	US-09-784-911-6
26	418	17.1	425	9	US-09-813-398-32
27	416.5	17.0	403	9	US-09-813-398-31
28	416.5	17.0	455	9	US-09-813-398-28
29	409.5	16.7	433	10	US-09-784-911-12
30	407.5	16.7	433	10	US-09-784-911-8
31	402.5	16.4	435	10	US-09-784-911-10
32	402.5	16.4	451	10	US-09-784-911-14
33	401.5	16.4	431	8	US-08-822-186-2
34	401.5	16.4	431	8	US-08-937-755-2
35	401.5	16.4	431	9	US-09-982-543A-10
36	401.5	16.4	431	10	US-09-045-331-2
37	401.5	16.4	431	10	US-09-828-607-2
38	400	16.3	72	9	US-09-945-182-13
39	399	16.3	432	10	US-09-361-741-3
40	372	15.2	427	12	US-10-002-278-2
41	359.5	14.7	143	12	US-10-002-278-3
42	357	14.6	72	9	US-09-945-182-15
43	355.5	14.5	356	12	US-10-084-037-1
44	354	14.5	118	9	US-09-859-211-37
45	354	14.5	118	9	US-09-880-708-15

ALIGNMENTS

RESULT 1
US-09-730-772-14
; Sequence 14, Application US/09730772
; Patent No. US2001001131A1
; GENERAL INFORMATION:
; APPLICANT: Luyten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; APPLICANT: Chang, Steven Chao-Huan
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/730,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,081
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartfeld, Neil S
; REGISTRATION NUMBER: 39,901
; REFERENCE/DOCKET NUMBER: NIH099.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 436 amino acids
; TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-730-772-14

Query Match 73.5%; Score 1798; DB 10; Length 436;
Best Local Similarity 80.8%; Pred. No. 2.6e-120;
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY 33 SSAELGSKGMSRKRGKMRAPRSDAGRE--GOEPOPQDEPR---AQOPRAQEP 86
DB 2 ASAEUGSAGKMTREGRMPRAPRENATAREPLDRQEPPEPQPEPPQPEAREPP 61
QY 87 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRQ 146
DB 62 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRQ 121
QY 147 KYLFVMSLKEELVGAELRFRQAPSPAGPLHVQFPCLSPLLLDARTLDPOG 206
DB 122 KYLFVSTLSKEELVGADVRLFRQAPALAPPAAALRLP-VAPAAAGSAEP-GPAG 179
QY 207 APPAGWEFVWQGRHOPKOLCLELRAAW-GELDAGEAEARARGPOPPDLSLGF 265
DB 180 APRGWEFVWQGRHOPKOLCLELRAAWGEGPAAEDEARTPGQPPPDLSLGF 239
QY 266 GRVRPPOERALLVFTSRQKNLFAEMREQLGSA-EAAGPGAGAGSW-----PPPS 317
DB 240 GRVRTPOERALLVFTSRQKTLFAEMREQLGSATEVVGPGGAGSGPPPPPPPPPS 299
QY 318 GAPDARWLPSPGRRRRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWIIAPLE 377
DB 300 GTPDAGLWSPSG-RRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWIIAPLE 358
QY 378 YEAYHCEGVCDPLRSHLEPTNHAIQTILMNSMDPGSTPPCCVPTKLTPI SILYIDAGN 437
DB 359 YEAYHCEGVCDPLRSHLEPTNHAIQTILMNSMDPGSTPPCCVPTKLTPI SILYIDAGN 418
QY 438 NVVYKQYEDMVYVSCGCR 455
DB 419 NVVYNEYEMVYVSCGCR 436

RESULT 2
US-09-735-849-14
Sequence 14, Application US/09735849
Patent No. US20010037017A1
GENERAL INFORMATION:
APPLICANT: Luyten, Frank P.
APPLICANT: Moos, Jr., Malcolm
APPLICANT: Chang, Steven Chao-Huan
TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09735.849
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836.081
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Bartfeld, Neil S
REGISTRATION NUMBER: 39,901
REFERENCE/DOCKET NUMBER: NIH099.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 436 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-735-849-14

Query Match 73.5%; Score 1798; DB 10; Length 436;
Best Local Similarity 80.8%; Pred. No. 2.6e-120;
Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY 33 SSAELGSKGMSRKRGKMRAPRSDAGRE--GOEPOPQDEPR---AQOPRAQEP 86
DB 2 ASAEUGSAGKMTREGRMPRAPRENATAREPLDRQEPPEPQPEPPQPEAREPP 61
QY 87 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRQ 146
DB 62 GRGPRVVPHEYMLSIYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDLSHTPLRQ 121
QY 147 KYLFVMSLKEELVGAELRFRQAPSPAGPLHVQFPCLSPLLLDARTLDPOG 206
DB 122 KYLFVSTLSKEELVGADVRLFRQAPALAPPAAALRLP-VAPAAAGSAEP-GPAG 179
QY 207 APPAGWEFVWQGRHOPKOLCLELRAAW-GELDAGEAEARARGPOPPDLSLGF 265
DB 180 APRGWEFVWQGRHOPKOLCLELRAAWGEGPAAEDEARTPGQPPPDLSLGF 239
QY 266 GRVRPPOERALLVFTSRQKNLFAEMREQLGSA-EAAGPGAGAGSW-----PPPS 317
DB 240 GRVRTPOERALLVFTSRQKTLFAEMREQLGSATEVVGPGGAGSGPPPPPPPPPS 299
QY 318 GAPDARWLPSPGRRRRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWIIAPLE 377
DB 300 GTPDAGLWSPSG-RRRTAFASRGKRGKKSRLRCSKKPLHVNFKELGWDWIIAPLE 358
QY 378 YEAYHCEGVCDPLRSHLEPTNHAIQTILMNSMDPGSTPPCCVPTKLTPI SILYIDAGN 437
DB 359 YEAYHCEGVCDPLRSHLEPTNHAIQTILMNSMDPGSTPPCCVPTKLTPI SILYIDAGN 418
QY 438 NVVYKQYEDMVYVSCGCR 455
DB 419 NVVYNEYEMVYVSCGCR 436

RESULT 3
US-09-945-182-26
Sequence 26, Application US/09945182
Patent No. US20020160494A1
GENERAL INFORMATION:
APPLICANT: Celeste, Anthony J.
APPLICANT: Wozney, John
APPLICANT: Rosen, Vicki A.
APPLICANT: Wolfman, Neil
APPLICANT: Thomsen, Gerald H.
APPLICANT: Melton, Douglas A.
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/945,182
FILING DATE: 31-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,324
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-945-182-26

Query Match 71.4%; Score 1747; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 7.6e-117;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 137 DLSHTPLRQKYLFDVSMLSKDEELVGAELRLFRQAPSPAGPAGPLHVQLPCLSPLL 196
Db 3 DLSHTPLRQKYLFDVSMLSKDEELVGAELRLFRQAPSPAGPAGPLHVQLPCLSPLL 62
QY 197 LDARTLDPQAGPAGVEFDVWGLRHQPWKQLCCLRAAWGELDAEAEARAGFPQPP 256
Db 63 LDARTLDPQAGPAGVEFDVWGLRHQPWKQLCCLRAAWGELDAEAEARAGFPQPP 122
QY 257 PDLRLSGFGRVRRPQERALLVFTSRQKLNFAEMRQLGSAEAGPAGAGSGWPPP 316
Db 123 PDLRLSGFGRVRRPQERALLVFTSRQKLNFAEMRQLGSAEAGPAGAGSGWPPP 182
QY 317 SGAPDARPLSPGRRRRRTAFASRHGKRSRCKKPLHVNFKELGWDWIIAPL 376
Db 183 SGAPDARPLSPGRRRRRTAFASRHGKRSRCKKPLHVNFKELGWDWIIAPL 242
QY 377 EYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPI SILYIDAG 436
Db 243 EYAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPI SILYIDAG 302
QY 437 NNVVYKQYEDMVVESCGR 455
Db 303 NNVVYKQYEDMVVESCGR 321

RESULT 4
US-09-945-182-32
Sequence 32, Application US/09945182
Patent No. US20020160494A1
GENERAL INFORMATION:
APPLICANT: Celeste, Anthony J.
Wozney, John
Rosen, Vicki A.
Wolfman, Neil
Thomsen, Gerald H.
Melton, Douglas A.
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive

CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/09/945,182
FILING DATE: 31-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,324
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-945-182-32

Query Match 49.1%; Score 1202.5; DB 9; Length 263;
Best Local Similarity 83.3%; Pred. No. 2.1e-78;
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;
QY 180 PAGPLHVQLPCLSPLLDARTLDPQAGPAGVEFDVWGLRHQPWKQLCCLRAAWGE 239
Db 7 PAG-----PTLRGSSGTQPR---PAG-KSFDVWGLRPPQWKQLCCLRAAWGE 51
QY 240 LDAGEAEARAGPQQPPDLRLSLGFGRRVRRPQERALLVFTSRQKLNFAEMRQLG 299
Db 52 LDAGDGTGARAGPQQPPDLRLSLGFGRRVRRPQERALLVFTSRQKLNFAEMRQLG 111
QY 300 AEAGPGAGAGSWPPSGAPDARPLSPGRRRRRTAFASRHGKRSRCKKPL 359
Db 112 AEA---AGAGSGCAPSGSPDTGSLPSPGRRRRRTAFASRHGKRSRCKKPL 167
QY 360 HVNFKELGWDWIIAPLEYEAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSC 419
Db 168 HVNFKELGWDWIIAPLEYEAYHCEGVCDPPLRSHLEPTNHAIIQTLNMSMDPGSTPPSC 227
QY 420 CVPTKLTPI SILYIDAGNNVVYKQYEDMVVESCGR 455
Db 228 CVPTKLTPI SILYIDAGNNVVYKQYEDMVVESCGR 263

RESULT 5
US-08-981-490B-1
Sequence 1, Application US/08981490B
Patent No. US20020045568A1
GENERAL INFORMATION:
APPLICANT: Hotten, Gertrud
APPLICANT: Pohl, Jens
APPLICANT: Bechtold, Rolf
APPLICANT: Paulista, Michael
APPLICANT: Unsicker, Klaus
TITLE OF INVENTION: USE OF MP52 OR MP121 FOR TREATING AND PREVENTING DISEASES OF T
FILE REFERENCE: 100564-07032
CURRENT APPLICATION NUMBER: US/08/981,490B
CURRENT FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: PCT/EP96/03065

; PRIOR FILING DATE: 1996-07-12
; PRIOR APPLICATION NUMBER: DE/195 25 416.3
; PRIOR FILING DATE: 1995-07-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-981-490B-1

Query Match 37.9%; Score 927.5; DB 8; Length 501;
Best Local Similarity 44.7%; Pred. No. 1.3e-58;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKGRSRKEGKMORAPRDSADAGREGQPPQPPQDEPRAQ-- 79
Db 81 GQTGGTQPKKDEPKLP-----PRPG-GPEPKPGHPPTQROTARTVTPKQLPGGKAP 134
QY 80 -----PRAQEPGRGRVVPHEYMLSIYRTYSIAEKLGINASFTQ 119
Db 135 PRAGSVSSFLKKAREPGPREPEKPEPPPTIPHEYMLSLYRTLSADRGKGNSSVKL 194
QY 120 SSKSANTITSFVDRGLDDLSHTPLRQKYLEFVSMDSKEELVGAEELRFRQAPS----- 174
Db 195 EAGLANITITSFIDKGQDDRGPV-VRKQRYVFDISAL-EKDGLLGAELRLIRLKKPSDTAKP 252
QY 175 -APWGPAGPLHVQLPCLLS-----PLLLDARTLDPGAPGAGVEFVQGLRH-QPWKQ 228
Db 253 APGGRGAQL--KLSSCSGGRQPSALLDVRV--PGLDGSQGEVFDIWKLFNFKNNSAQ 308
QY 229 LCLELRAANGELDAGAEARAGPQPPPPDLKSLGFRGRVRPPQBRALLVVFTRSORKN 288
Db 309 LCLELE-AW---ERGRA-----VDLRLGLGFDRAARQVHEKALFLVFGRTKKRD 352
QY 289 L-FAEMREQLGSAEAGAGGAGEGSWPPSPGADARPWLPSGRRRRRTAFASRHGRHG 347
Db 353 LFFNEIKARSGQDDKT-----VVEYLFQ--RRKRRAPLATRQGRKPS 393
QY 348 KKSRLCSKKPLHVNFKELGWDWIITAPLEYAHCEGVCDPFLRSHLEPTNHAIQTLM 407
Db 394 KNLKARCSKKALHVNFKMGWDWIITAPLEYEAFHCEGLCEFFLRSHLEPTNHAIQTLM 453
QY 408 NSMDPSTPPSCVPTKLPISILYIDAGNNVYKYQYEDMVVESCGR 455
Db 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVYKYQYEDMVVESCGR 501

RESULT 6

US-09-880-708-10
; Sequence 10, Application US/09880708
; Patent No. US20020165361A1
; GENERAL INFORMATION:

APPLICANT: Lee, Se-Jin
Huynh, Thanh
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:

ADDRESSEE: Gray Cary Ware & Freidenrich LLP
STREET: 4365 Executive Drive, Suite 1600
CITY: San Diego
STATE: CA

COUNTRY: USA
ZIP: 92121-2189

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/880,708

FILING DATE: 12-Jun-2001

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/145,060
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/003,144
; FILING DATE: 12-JAN-1993
; ATTORNEY/AGENT INFORMATION:
NAME: Lisa A. Haile, Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/057002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 858/677-1456
TELEFAX: 619/677-1465

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 495 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-880-708-10

Query Match 37.6%; Score 919.5; DB 9; Length 495;
Best Local Similarity 43.5%; Pred. No. 4.6e-58;
Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PGFQQAASISSSCSSABLGSTKGRSRKEGKMORAPRDSADAGREGQPPQPPQDE----- 74
Db 60 PGHGYGVATNARAK-GSSGQTQAKKDEPRKMPPRSG-----GSETKPGPSSQTQAAA 113
QY 75 -----PRAQOP-----RAQEP-----PGRGRVVPHEYMLSIYRT 104
Db 114 RTWTPKQQLPGGKASSKAGSAPSSFLKKKTREPTREPKEPPRPPIITPHEYMLSIYRT 173
QY 105 YSIAEKLGINASFFQSSKSANTITSFVDRGLDDLSHTPLRQKYLEFVSMDSKEELVGA 164
Db 174 LSDADRKGNSSVYLEAGLANITITSFIDKGQDDRG-PAVRKQRYVFDISAL-EKDGLLGA 231
QY 165 ELRLFRQAP--SAPMGPPAGPL-HVQLPCLLS-----PLLLDARTLDPGAPGAGVEFD 216
Db 232 ELRLRKKPLDVAKPAVPSGGRVAQLKLSGSPGRPAALLDVRV--PGLDGSQGEVFD 289
QY 217 VMOGLRH-QPWKQLCLELRAANGELDAGAEARAGPQPPPPDLKSLGFRGRVRPPQER 275
Db 290 IWKLFNFKNNSAQLCLELE-AW---ERGRA-----VDLRLGLGFERTARQVHEK 333
QY 276 ALLVVFTRSORKNL-FAEMREQLGSAEAGAGGAGSGWPPSPGADARPWLPSGRRRR 334
Db 334 ALFLVFGRTKKRDLFFNEIKARSGQDDKT-----VVEYLFQ--RRKR 374
QY 335 RTAFASRHRGKRRHKKSRRLKRLHVNFKELGWDWIITAPLEYAHCEGVCDPFLRSH 394
Db 375 RAPLANRQGRKPSKLNKARCSRKALHVNFKMGWDWIITAPLEYEAFHCEGLCEFFLRSH 434
QY 395 LEPTNHAIQTLMNSMDPGSTPPSCVPTKLPISILYIDAGNNVYKYQYEDMVVESCGR 454
Db 435 LEPTNHAIQTLMNSMDPESTPTCCVPTRLSPISILFIDSANNVYKYQYEDMVVESCGR 494
QY 455 R 455
Db 495 R 495

RESULT 7

US-09-813-398-37
; Sequence 37, Application US/09813398
; Patent No. US20020169292A1
; GENERAL INFORMATION:

APPLICANT: Bruce D. Weintraub

APPLICANT: Mariusz W. Szkudlinski

APPLICANT: University of Maryland

TITLE OF INVENTION: CYSTINE KNOT GROWTH FACTOR MUTANTS

FILE REFERENCE: UOFMD.003C1

CURRENT APPLICATION NUMBER: US/09/813,398

```

; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: PCT/US99/05908
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/US98/19772
; PRIOR FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 502
; TYPE: PRT
; ORGANISM: HOMO SAPIEN
US-09-813-398-37

```

TELEPHONE: 019-235-8300
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 13:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-730-772-13

Query Match 36.9%; Score 903.5; DB 10; Length 501;
Best Local Similarity 44.9%; Pred. No. 6.3e-57;
Matches 207; Conservative 63; Mismatches 108; Indels 83; Gaps

QY 38 GSTGMRSRKEGKMQRAP--RDSADAGEQEPQPRQD-----EPRAQPP----- 80
| | | | | : : : | | | | | : | |
Db 81 GQTGGTLPQPKDEPKPLPRPGGPEPKRGHPPTQRTQATARTVTPGQLPGGKAPKAGSV 140
| | | | | : | | | | | : | | | | | : | |

QY 81 -----RAQEP-----PGRGPWVHEYMLSYRTYSIAEKLGINASFQSSKAN 125
| | | | | : | | | | | : | | | | | : | |
Db 141 PSSFLKKAREGPPPREKPEFPFPPIITHEYMLSYRLTSDADRGKGNSSVKLEAGLAN 200
| | | | | : | | | | | : | | | | | : | |

QY 126 TITSFVDRGLDLSHTPLRQKYLFDVMSLSKEELVGAELRLFQAPSAPMGPPAGP-- 183
| | | | | : | | | | | : | | | | | : | |
Db 201 TITSFTDKGQDDRGVP-VRKQRYVDFISAL-EKDGILGALIELRLRKPS-DTAKPAVPRS 257
| | | | | : | | | | | : | | | | | : | |

QY 184 ---LHVQLFPCLs----PLLLDARTLDPGAPPAGWEVDFWQGLRH-OPWQQLCLLELA 235
| | | | | : | | | | | : | | | | | : | |
Db 258 RRAAQLKSLSCPSGRQPAALLDVRSV--PGLDSGWEVDFIWKLFNFNSAQLCLLEL-- 313
| | | | | : | | | | | : | | | | | : | |

QY 236 AWGELDAGEAEARAGQPPPPPDRLSLGFGRRVRPPQPERALLVWTRTSQRKNL-FAEMR 294
| | | | | : | | | | | : | | | | | : | |
Db 314 -----EAWERGR-----TVDLRGLGLFDRAARQVHEKALFLVFGRTKKRDLFFNEIK 359
| | | | | : | | | | | : | | | | | : | |

```

```

Db      360  ARSGQDDKT-----VVEYLFQ--RRKRAPSATRC
QY      355  SKRPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPLSHLEPTNHA
Db      401  SRKALHVNFKDGMGDWIIAPLEYEAFGCEGCEPPLRSHLEPTNHA
QY      415  TTPSCCVPTKLTPISTILYIDAGNNVVKQYEDMWVYVSGCR 455
Db      461  TPTCCVPTPLSPISILFLFIDSANNVVKQYEDMWVYVSGCR 501

RESULT 9
US-09-735-849-13
; Sequence 13, Application US/09735849
; Patent No. US20010037017A1
; GENERAL INFORMATION:
; APPLICANT: Luyten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; APPLICANT: Chang, Steven Chao-Huan
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:

```

```

RESULT 9
US-09-735-849-13
; Sequence 13, Application US/09735849
; Patent No. US20010037017A1
; GENERAL INFORMATION:
; APPLICANT: Layten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; APPLICANT: Chang, Steven Chao-Huan
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
;

```

ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/735,849
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/836,081
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bartfeld, Neil S
REGISTRATION NUMBER: 39,901
REFERENCE/DOCKET NUMBER: NIH099.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-735-849-13

Query Match 36.9%; Score 903.5; DB 10; Length 501;
Best Local Similarity 44.9%; Pred. No. 6.3e-57;
Matches 207; Conservative 63; Mismatches 108; Indels 83; Gaps 16;
QY 38 GSTKGRSRKEGKMORAP--RSDAGREGQEPQPRQD---EPRAQQP----- 80
Db 81 QGTGGLTQPKDEPKLPFGPGPEPKRHPPTQRTATVTTPKGLPGGKAPPKAGSV 140
QY 81 -----RAQEP-----PGRGPRVVPHEYMLSYRTYSIAEKLGINASFQSKSAN 125
Db 141 PSSFLKKAREPGPREPEPRPPITTHEYMLSLYRTLSDADRKGGNSVKLEAGLAN 200
QY 126 TTTSFVDRGLDLSHTPLRRQYLFVDSMLSKEELVGAELRLFRQAPSPAGPPAGP-- 183
Db 201 TTTSFDKGQDGRGPV-VRKQRYVFDISAL-EKDGLLGAELRLRKPS-DTAKPAVPRS 257
QY 184 ---LHVQLFPCLs---PILLDARTLDPOGAPPAGWEVDVMOGLRH-QPWKOLCLELRA 235
Db 258 RRAAQKLLSCSPSGQPAALLDVRV--PGLDGSQGEVFDIWKLFNFKNASQOLCLEL-- 313
QY 236 AMGELDAGAEARARGPQPPPDRLSLRGLFGRVRPPQERALLVVFTRSORKNL-FAEMR 294
Db 314 -----EAWERGR-----TVDLRGLGFDRAARQVHEKALFLVFGRTKKDLFENEIK 359
QY 295 EQLGSAEAGPAGAGSGWPPSGADAPRWLPSPGRRRRRTAFASRRHGKRSRLRC 354
Db 360 ARSGQDDKT-----VVEYLFQ--RRKRRAPSATROCKRPKSKNLKARC 400
QY 355 SKKPLHVNFKELGDDWIIAPLEYEAYHCEGVCDPPLRSHLEPTNHAIQTLLMNSMDPS 414
Db 401 SKKALHVNFKDGMGDDWIIAPLEYEAFGCEPPLRSHLEPTNHAVIQTLLMNSMDPS 460
QY 415 TTPSCVPTKLTPIISILYIDAGNNVYKQYEDMVVESCGR 455
Db 461 TPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESCGR 501

RESULT 10

US-09-945-182-34
Sequence 34, Application US/09945182
Patent No. US2002016049A1
GENERAL INFORMATION:
APPLICANT: Celeste, Anthony J.
Wozney, John
Rosen, Vicki A.
Wolfman, Neil
Thomsen, Gerald H.
Melton, Douglas A.
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/945,182
FILING DATE: 31-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,324
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-945-182-34

Query Match 31.3%; Score 765; DB 9; Length 388;
Best Local Similarity 51.2%; Pred. No. 3e-47;
Matches 172; Conservative 36; Mismatches 78; Indels 50; Gaps 10;
QY 146 QKYLEFVDSMLSKEELVGAELRLFRQ-APSAWGPAGPLHVQLFPCL-----LSPILLDAR 200
Db 77 QSFLEFVDSLSNDADEVVGAELRVLRGSPESGPGSWTSPPLLLSLTCTPGAARAPRLLYSR 136
QY 201 TLDPOCAPAGWEVDVMOGLRH-----OPWKOLCLELRAAMGELDAGAEARARGPQOP 255
Db 137 AAEP--LVQQRWEAFVADAMRRHRREPFRPRAFCULLLRVAVGPV-----P 180
QY 256 PPPDLRLSLGF---GRRVRPPQERALLVVFTRSORK-NLFAEMREQ---LGSAAEAGPGAG 308
Db 181 SPLALRRLGFGWGGGSAEERAVLVVSSRTQRKESLREIRAQAARALGAALASEP--- 237
QY 309 AEGSWPPSGADAPRWLPSPGRRRRRTAFASRH-----GKRHGKRSRLRSKKPL 359
Db 238 -----LPDPGTASPRAVIGGRRRRRTALAGTRTAQSGGGAGRGHGRGRRSGRSRKPL 292
QY 360 HVNFKELGDDWIIAPLEYEAYHCEGVCDPPLRSHLEPTNHAIQTLLMNSMDPGSTPPSC 419
Db 293 HVDFKELGDDWIIAPLEYEAYHCEGLCDFPLRSHLEPTNHAIQTLLNSMAPDAPASC 352
QY 420 CVPTKLTPIISILYIDAGNNVYKQYEDMVVESCGR 455
||| :|:||||||| |||||||||||:|||||

[illegible]

Db 243 HVRISRLHQDEHSWSQIRPLLVTFGHDKGHPHLHREKRTALAGTRTAQGGGGA----- 298

QY 314 PPSGAPDARPWLPSPGRRRRRTAFASRHGKRGKSRKRCCKPLHVNFKELGWDWII 373

Db 299 -----GRHGRGRGRCSRKPLHVDKELGWDWII 329

QY 374 APLEYAYHCEGVCDPFLRSHLEPTNHAIIQTLNMSMDPGSTPPSCCVPTKLTPIISILYI 433

Db 330 APLDYAYHCEGLCDPFLRSHLEPTNHAIIQTLNMSMDPAAPASCVCVPARLSPISILYI 389

QY 434 DAGNNVYKQYEDMVVESCGR 455

Db 390 DAANNVYKQYEDMVVESCGR 411

RESULT 13

US-09-945-182-30

; Sequence 30, Application US/09945182

; Patent No. US20020160494A1

; GENERAL INFORMATION:

; APPLICANT: Celeste, Anthony J.

; Wozney, John

; Rosen, Vicki A.

; Wolfman, Neil

; Thomsen, Gerald H.

; Melton, Douglas A.

; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS

; NUMBER OF SEQUENCES: 35

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENETICS INSTITUTE, INC.

; STREET: 87 Cambridgepark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/945,182

FILING DATE: 31-Aug-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/808,324

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Lazar, Steven R.

REGISTRATION NUMBER: 32,618

REFERENCE/DOCKET NUMBER: 5202-D

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617 498-8260

TELEFAX: 617 876-5851

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 240 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 30:

US-09-945-182-30

Query Match 25.3%; Score 618; DB 9; Length 240;

Best Local Similarity 55.6%; Pred.No.4.5e-37;

Matches 130; Conservative 20; Mismatches 34; Indels 50; Gaps 7;

QY 260 LRSGLGFRVRVP-----QERALLVFTSRQK-NLFAEMREQLSAEAGPGAG 308

Db 19 LRLGFGX----PGGDGGGTAXEERALLVISRTRQKESLFEIRAQARLRAA----- 69

QY 309 AEGSWPPSGADPARWLPSPGRRRRRTAFASRH----- 342

Db 70 AE---PPDPGPGAGSRKANLGRRRQRTALAGTRGXGSGGGGGGGGGGGGGG 126

QY 343 -GRRHGKSRKRCCKPLHVNFKELGWDWIIAPLEYAYHCEGVCDPFLRSHLEPTNHA 401

Db 127 ARGHGRGRGRCSRKSLHVDKELGWDWIIAPLDYAYHCEGVCDPFLRSHLEPTNHA 186

QY 402 IIQTLNMSMDPGSTPPSCCVPTKLTPIISILYIDAGNNVYKQYEDMVVESCGR 455

Db 187 IIQTLNMSMDPAAPASCVCVPARLSPISILYIDANNVYKQYEDMVVESCGR 240

RESULT 14

US-08-945-459A-1

; Sequence 1, Application US/08945459A

; Patent No. US20020102633A1

; GENERAL INFORMATION:

; APPLICANT: MAKISHIMA, FUSAO; TAKAMATSU,

; APPLICANT: HIRIYUKI, MIKI, HIDEO; KAWAI,

; APPLICANT: SHINJI, KIMURA, MICHIO; MATSUMOTO,

; APPLICANT: TOMOAKI; KATSUURA, MIEKO; ENOMOTO,

; APPLICANT: KOICHI; SATOH, YUSUKE

; TITLE OF INVENTION: A NOVEL PROTEIN AND

; TITLE OF INVENTION: PROCESS FOR PREPARING THE SAME

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS

; ADDRESSEE: LLP

; STREET: 600 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10016

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/945,459A

FILING DATE: 09-DEC-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/01062

FILING DATE: 19-APR-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP7/322403

FILING DATE: 17-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP7/93664

FILING DATE: 19-APR-1995

ATTORNEY/AGENT INFORMATION:

NAME: CHARLES A. MUSERLIAN

REGISTRATION NUMBER: 19,683

REFERENCE/DOCKET NUMBER: 146.1275

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 661-8000

TELEFAX: (212) 661-8002

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PEPTIDE

ORIGINAL SOURCE:

ORGANISM: HOMOSAPIENS

TISSUE TYPE: FETUS

FEATURE:

NAME/KEY: MP52

LOCATION: 383 TO 501

US-08-945-459A-1

Query Match 22.4%; Score 549; DB 8; Length 119;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:50:50 ; Search time 91 Seconds
(without alignments)
1030.236 Million cell updates/sec

Title: US-09-825-751A-20
Perfect score: 2447
Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1253	51.2	399	13 Q9W753	Q9W753 xenopus lae
2	1091.5	44.6	412	13 Q12938	Q12938 brachydanio
3	917	37.5	413	13 Q9DGN4	Q9DGN4 xenopus lae
4	902	36.9	500	13 Q9W6G0	Q9W6G0 gallus gall
5	870	35.6	447	6 Q9BDW8	Q9BDW8 cercopithec
6	841.5	34.4	324	13 Q9YHW9	Q9YHW9 gallus gall
7	737	30.1	294	6 Q9BDW9	Q9BDW9 macaca fasc
8	705	28.8	441	11 Q99MY1	Q99MY1 mus musculus
9	698.5	28.5	261	13 Q9W6C0	Q9W6C0 brachydanio
10	664.5	27.2	257	13 Q42303	Q42303 brachydanio
11	612.5	25.0	126	13 Q93573	Q93573 gallus gall
12	507	20.7	417	5 Q9XVQ7	Q9XVQ7 lytechinus
13	457.5	18.7	391	13 Q8UVQ2	Q8UVQ2 brachydanio
14	456.5	18.7	391	13 Q8UVQ8	Q8UVQ8 brachydanio
15	453	18.5	398	13 Q90YD7	Q90YD7 xenopus tro
16	451	18.4	277	13 Q90Y82	Q90Y82 lampetra ja

17	448.5	18.3	391	13 Q90YJ3	Q90YJ3 brachydanio
18	445	18.2	400	13 Q57574	Q57574 brachydanio
19	442	18.1	364	13 Q9PVK1	Q9PVK1 gallus gall
20	439.5	18.0	405	5 Q9U5E8	Q9U5E8 ptychodera
21	438	17.9	361	5 Q96504	Q96504 branchiosteo
22	438	17.9	400	13 Q13107	Q13107 brachydanio
23	438	17.9	411	5 Q9U418	Q9U418 branchiosteo
24	431.5	17.6	390	13 Q91597	Q91597 xenopus lae
25	428.5	17.5	509	5 Q8WS99	Q8WS99 archaster t
26	426	17.4	289	5 Q9XVQ8	Q9XVQ8 strongyloce
27	424.5	17.3	386	13 Q13109	Q13109 brachydanio
28	424	17.3	337	6 Q9MZV5	Q9MZV5 canis famil
29	412	16.8	614	5 Q91720	Q91720 drosophila
30	409.5	16.7	400	13 Q73818	Q73818 xenopus lae
31	409.5	16.7	411	13 Q93369	Q93369 brachydanio
32	408.5	16.7	400	13 Q91703	Q91703 xenopus lae
33	405.5	16.6	453	13 Q87373	Q87373 gallus gall
34	404.5	16.5	400	13 Q90YD6	Q90YD6 xenopus tro
35	403.5	16.5	411	13 Q57573	Q57573 brachydanio
36	402	16.4	301	5 Q97390	Q97390 crassostrea
37	401.5	16.4	204	5 Q9XZ69	Q9XZ69 tripneustes
38	398	16.3	391	13 Q87380	Q87380 brachydanio
39	397.5	16.2	432	13 Q9PTP9	Q9PTP9 brachydanio
40	396.5	16.2	398	13 Q918T6	Q918T6 gallus gall
41	395.5	16.2	588	5 Q9VQC6	Q9VQC6 drosophila
42	391	16.0	313	13 Q91403	Q91403 gallus gall
43	390.5	16.0	128	5 Q95W38	Q95W38 schistocerc
44	390	15.9	443	5 Q76851	Q76851 halocynthia
45	385.5	15.8	182	13 Q90Y81	Q90Y81 lampetra ja

ALIGNMENTS

RESULT 1

Q9W753 Q9W753 PRELIMINARY; PRT; 399 AA.
ID AC Q9W753;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Growth and differentiation factor 6.
GN GDF6.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99396700; PubMed=10393114;
RA Chang C., Hemmati-Briylanlou A.;
RT "Xenopus GDF6, a new antagonist of noggin and a partner of BMPs."
RL Development 126:3347-3357(1999).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF155125; RAD38402.1; -.
DR HSSP; PI2643; 3BMP.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
SQ SEQUENCE 399 AA; 45571 MW; C549D973B50B8517 CRC64;

Query Match 51.2%; Score 1253; DB 13; Length 399;
Best Local Similarity 58.8%; Pred. No. 1.7e-92;
Matches 270; Conservative 40; Mismatches 85; Indels 64; Gaps 9;

QY 17 LWDLPGFQQASISSSSSAELGKMGMSRKEGKMRAPRDSAGREGQEPQPRQDEPR 76
DB 13 LWSAWLWAAAAAGLLSSRLRLTGGAAR-RAIATLHGHPKD-----52
QY 77 AQOPRAQEPGPRVVPHEYMLSIYRTSYIAEKLGINASFFOSSKSANTITSFVDRGLD 136
DB 53 -HLLLSKEPLGKG-TVPHEYMLSLYKSJSSABHQDSRSATSRNRSTNTITSFVDRGHD 110
QY 137 DLSHTPLRROKYLFDVSMLSKDEELVGAELRLFRQAP--SAPWGPAGPLHVLQFPC---191
DB 111 DLPQP-PEQKFLFDRSLPERDEIVGAELRIILRRPEDTSSSESGLVHLNLYCTPTG 169
QY 192 -LSPLLDARTLDOGAPPAGNVEFDWQ--GLRHQFWKQLCLELRAANGELDAGEAAR 248
DB 170 WQEPKLTIDSPADLLDTVFSKVEVFNWKAAGNRRLSGETLCFMLKI-----VSDIKAT 223
QY 249 ARGQPOPPDLRLSLGFRVRPQERALLVVTTRSQK-NLFAEMREQLGSAAGPGA 307
DB 224 AFEFVQ-----IGFSRGGQLPHEKALLVSSHSKRRLENKEIKSI-----G 269
QY 308 GAGSWPPSPGADPAPLWPSGRRRRRTAFASR--HGRKHGKSLRCSKKPLHVNKE 365
DB 270 NPKEPSPGQGSIAK-----RRKRTLPTRTNGKGHAKSKTRCSKKPLLVNKE 323
QY 366 LGWDDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKL 425
DB 324 LGWDDWIIAPLDYAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPESTPPSCCVPSKL 383
QY 426 TPISILYIDAGNNVYKQYEDMVVESCGR 455
DB 384 SPISILYIDSGNNVYKQYEDMVVESCGR 413

RESULT 4
Q9W6G0 ID Q9W6G0 PRELIMINARY; PRT; 500 AA.
AC Q9W6G0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE GDF-5 protein.
GN GDF5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99146893; PubMed=10021348;
RA Francis-West P.H., Abdelfattah A., Chen P., Allen C., Parish J.,
RA Ladhier R., Allen S., Macpherson S., Luyten F.P., Archer C.W.;
RT "Mechanisms of GDF-5 action during skeletal development.";
RL Development 126:1305-1315(1999).
CC -I- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF123389; RAD30451.1; -.
DR HSSP; PI2643; 3BMP.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
SQ SEQUENCE 500 AA; 55952 MW; 1DE8385A3119A598 CRC64;

Query Match 36.9%; Score 902; DB 13; Length 500;
Best Local Similarity 44.2%; Pred. No. 3.2e-64;
Matches 204; Conservative 69; Mismatches 129; Indels 60; Gaps 14;

QY 2 DTPRVLLSAVFLISLWDLPGFQQASISSSSSAELGKMGMSRKEGKMRAPRDSAG 61
DB 91 ESQVLSRTAVTEAKVGHLPSPRSVAVRTYKVPNLGSKAAAKTKTG-----G 139
QY 62 REQEPQPRQDEPRQAQPPGPRVVPHEYMLSIYRTSYIAEKLGINASFFQSS 121
DB 140 TDASSFKTKTKKEPVQR-EAKETFSR-PPITPHEYMLSIYRTSLSDAERKGVNGSKLEA 197
QY 122 KSANTITSFVDRGLDLSHTPLRROKYLFDVSMLSKDEELVGAELRLFRQAPSPAPWCP 181
DB 198 GLANTITSFIDKQDERAPT-IRKQYIIFDISAL-EKDGLLVAEILRLKNPNDTWNKSHS 255
QY 182 G--PLHVQLFPC-----LSPLLLDARTLDOGAPPAGNVEFDWQGLRH-OPWQLQCLCLR 234
DB 256 SGKTSQVKLFSCSTNQATLLDSRTVSTIDTPK--WEVFDIWKLFNFKNLVNLCFELE 313
QY 235 AANGELDAGEAARARQPOPPDLRLSLGFRVRPQERALLVVTTRSQK-NLFAEM 293
DB 314 T-----FDRGRA-----VDLRTVGFNRTGRQVNEKALFLVFGRTKKRDLFNEI 357
QY 294 REQLGSAEAGPCAGAGSWPPSPGADPAPLWPSGRRRRRTAFASRHKRHKKSLR 353
DB 358 KARSGDDDKT-----VYELFNQ--RRRRRAPLATROGKRPSKNLKP 398
QY 354 CSKKPLRVHFKELGWDDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPG 413
DB 399 CSKRALHVFNKDGMDDWIIAPLEYAYHCEGLCEPLRSHLEPTNHAIQTLMNSMDPE 458
QY 414 STPPSCCVPTKLTPISILYIDAGNNVYKQYEDMVVESCGR 455
DB 459 STPPSCCVPTKLTPISILYIDAGNNVYKQYEDMVVESCGR 500

RESULT 5
Q9BDW8 ID Q9BDW8 PRELIMINARY; PRT; 447 AA.
AC Q9BDW8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Growth/differentiation factor 7.
GN GDF7.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21136583; PubMed=11238730;
RA Watakabe A., Fujita H., Hayashi M., Yamamori T.;
RT "Growth/differentiation factor 7 is preferentially expressed in the
RT primary motor area of the monkey neocortex.";
RL J. Neurochem. 76:1455-1464(2001).
CC -I- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF254569; AAK30842.1; -.
DR HSSP; PI2643; 3BMP.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
SQ SEQUENCE 447 AA; 46866 MW; DF46D591925A8391 CRC64;

Query Match 35.6%; Score 870; DB 6; Length 447;
Best Local Similarity 49.4%; Pred. No. 1e-61;
Matches 201; Conservative 40; Mismatches 114; Indels 52; Gaps 12;


```
Db 58 VSDSNNEAVHFGM-----LGLSREDQOQTHERRALLVAFSQARRKENLFREIREKIRAM 109
Qy 301 EAAGPGAGAGSWPPSPGAPDAPWLPSPGRRRTAFASR-----HGKRGKSKRLR 353
Db 110 KS-----RKFSNPTPEHSIKGHP-----RHRRTALAGRPGVGPITSGGKGGRRTR 159
Qy 354 CSKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPG 413
Db 160 CSKPLHVNFKELGWDWIIAPLDYAYHCEGLCDFPLRSHLEPTNHAIQTLMNSMDPE 219
Qy 414 STPPSCCVPTKLTPTISILYIDAGNNVVKQYEDMVVESCGR 455
Db 220 STPPSCCVPSKLSILYIDSGNNVVKQYEDMVVESCGR 261

RESULT 10
Q42303 ID 042303 PRELIMINARY; PRT; 257 AA.
AC 042303;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE Contact (Fragment).
GN GDF5.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97398455; PubMed=9256353;
RA Bruneau S., Mourrain P., Rosa F.M.;
RT "Expression of contact, a new zebrafish DVR member, marks mesenchymal
RT cell lineages in the developing pectoral fins and head and is
RT regulated by retinoic acid.";
RL Mech. Dev. 65:163-173(1997).
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; Y12005; CAA72733.1; -.
DR HSSP; P12643; 3BMP.
DR ZFIN; ZDB-GENE-990415-39; gdf5.
DR InterPro; IPR002400; GF_cysknot.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD000357; TGFb; 1.
DR SMART; SM00284; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
FT NON_TER 1
FT CHAIN 140 257
SQ SEQUENCE 257 AA; 29787 MW; 6D64F0542F948849 CRC64;

Query Match 27.2%; Score 664.5; DB 13; Length 257;
Best Local Similarity 51.8%; Pred. No. 1.6e-45;
Matches 144; Conservative 32; Mismatches 53; Indels 49; Gaps 9;

Qy 186 VOLFPCLs-----PLLLDARTLDPOGAPGAGWVFVDW---QGLRHQPKQLCLELAAG 238
Db 21 LRLFTCASGKNAVLQARPDFSHA--SYEVFIDWVKFRNRP--QJCLELDA--- 73
Qy 239 ELDAEAEARAGPOQPPLDLRSIFGRRVRPPOERALLVVFTSRQKNLF-ADMRRQL 297
Db 74 -VDHGR-----PLDLRLGLSRAGQTKFAFFVFGTKRGLFYNIKAR- 119
Qy 298 GSAEAGPGAGAGSWPPSPGAPDAPWLPSPGRRRTAFASRHKRGKSKRLRCSKK 357
Db 120 -----SCHDNKTVVEYLFQTRMRRAPLR-GKPIKNPKQRCNRK 159
Qy 358 PLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAIQTLMNSMDPGSTPP 417
Db 160 QHVNFKELGWDWIIAPLEYAFHCDGVCDFPIRSHLEPTNHAIQTLMNSMDPRSTPP 219
```

```
Qy 418 SCCVPTKLTPTISILYIDAGNNVVKQYEDMVVESCGR 455
Db 220 TCCVPTRLSPISILYIDSANNVVKQYEDMVVESCGR 257

RESULT 11
O93573 ID 093573 PRELIMINARY; PRT; 126 AA.
AC 093573;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative growth/differentiation factor 6/7 (Fragment).
GN GDF6/7.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99026113; PubMed=9808626;
RA Lee K.J., Mendelsohn M., Jessell T.M.;
RT "Neuronal patterning by BMPs: a requirement for GDF7 in the generation
RT of a discrete class of commissural interneurons in the mouse spinal
RT cord.";
RL Genes Dev. 12:3394-3407(1998).
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; AF089086; AAC97113.1; -.
DR HSSP; P12643; 3BMP.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001839; TGFb.
DR Pfam; PF00019; TGF-beta; 1.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Glycoprotein.
FT NON_TER 1
SQ SEQUENCE 126 AA; 14265 MW; CB824D280F44A394 CRC64;

Query Match 25.0%; Score 612.5; DB 13; Length 126;
Best Local Similarity 86.5%; Pred. No. 1e-41;
Matches 109; Conservative 10; Mismatches 6; Indels 1; Gaps 1;

Qy 331 RRRRTAFASRHGKR-HGKSKRLRCSKKPLHVNFKELGWDWIIAPLEYAYHCEGVCDF 389
Db 1 RKRRTTTAARSGRGHGKAKTRCSRKPLHVNFKELGWDWIIAPLDYAYHCEGVCDF 60
Qy 390 PLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPTISILYIDAGNNVVKQYEDMVV 449
Db 61 PLRSHLEPTNHAIQTLMNSMDPESTPPSCCVPSKLSILYIDSGNNVVKQYEDMVV 120
Qy 450 ESCGCR 455
Db 121 ETCGCR 126

RESULT 12
O9XY07 ID 09XY07 PRELIMINARY; PRT; 417 AA.
AC 09XY07;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Bone morphogenetic protein BMP2/4.
GN BMP2/4.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:33:40 ; Search time 37 Seconds
(without alignments)
510.047 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESQGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	1798	73.5	436	1	GDF6_BOVIN	P55106 bos taurus
2	927.5	37.9	501	1	GDF5_HUMAN	P43026 homo sapien
3	918.5	37.6	495	1	GDF5_MOUSE	P43027 mus musculus
4	691	28.2	125	1	GDF6_MOUSE	P43028 mus musculus
5	549	22.4	151	1	GDF7_MOUSE	P43029 mus musculus
6	471	19.2	398	1	BMP4_XENLA	P25703 xenopus lae
7	464	19.0	398	1	BMP8_XENLA	P30884 xenopus lae
8	453	18.5	405	1	BMP4_CHICK	Q90752 gallus gall
9	449.5	18.4	395	1	BMP2_RABIT	O45564 oryctolagus
10	441.5	18.0	408	1	BMP4_HUMAN	P12644 homo sapien
11	436.5	17.8	408	1	BMP4_MOUSE	P21275 mus musculus
12	434	17.7	393	1	BMP2_RAT	P49001 rattus norv
13	432.5	17.7	408	1	BMP4_RAT	Q06826 rattus norv
14	431.5	17.6	396	1	BMP2_HUMAN	P12643 homo sapien
15	430	17.6	409	1	BMP4_RABIT	O45576 oryctolagus
16	429.5	17.6	461	1	DVRL1_STRPU	P48969 strongyloce
17	427.5	17.5	396	1	BMP2_DAMDA	O19006 dama dama
18	423	17.3	394	1	BMP2_MOUSE	P21274 mus musculus
19	419.5	17.1	408	1	BMP4_DAMDA	Q25607 dama dama
20	418.5	17.1	452	1	BMP5_MOUSE	P49003 mus musculus
21	418	17.1	424	1	BMP10_HUMAN	O95393 homo sapien
22	417.5	17.0	353	1	BMP2_CHICK	Q90751 gallus gall
23	416.5	17.0	402	1	BMP8_HUMAN	P34820 homo sapien
24	416.5	17.0	454	1	BMP5_HUMAN	P22003 homo sapien
25	413.5	16.9	401	1	BMP4_XENLA	P30885 xenopus lae
26	412.5	16.9	399	1	BMP4_MOUSE	P34821 mus musculus
27	407	16.6	426	1	BMP7_XENLA	P30886 xenopus lae
28	406	16.6	420	1	BMP10_XENLA	O95229 mus musculus
29	403	16.5	621	1	DECA_DROPS	P91699 drosophila
30	401.5	16.4	431	1	BMP7_HUMAN	P18075 homo sapien
31	400	16.3	430	1	BMP7_MOUSE	P23359 mus musculus
32	398	16.3	428	1	GDF2_MOUSE	Q9WV56 mus musculus
33	395.5	16.2	588	1	DECA_DROME	P07713 drosophila

34	395.5	16.2	593	1	DECA_DROSI	P91706 drosophila
35	394.5	16.1	510	1	BMP6_MOUSE	P20722 mus musculus
36	394.5	16.1	513	1	BMP6_HUMAN	P22004 homo sapien
37	388.5	15.9	395	1	UNIV_STRPU	P48970 strongyloce
38	388.5	15.9	429	1	GDF2_HUMAN	Q9UK05 homo sapien
39	380.5	15.5	372	1	DECA_TRICA	Q26974 tribolium c
40	377	15.4	354	1	NODA_MOUSE	P43021 mus musculus
41	372	15.2	427	1	DSL1_CHICK	P34822 gallus gall
42	358.5	14.7	399	1	BM8B_MOUSE	P55105 mus musculus
43	351	14.3	476	1	BM3B_RAT	P55107 mus musculus
44	350.5	14.3	478	1	BM3B_HUMAN	P55108 homo sapien
45	344.5	14.1	355	1	DVRL1_BRARE	P35621 brachydanio

ALIGNMENTS

RESULT 1
GDF6_BOVIN STANDARD; PRT; 436 AA.
AC P55106;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth/differentiation factor 6 precursor (GDF-6) (Cartilage-derived
DE morphogenetic protein 2) (CDMP-2) (Fragment).
GN GDF6 OR CDMP2
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular cartilage;
RX MEDLINE=95050604; PubMed=7961761;
RA Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,
RYba N.J.P., Kozak C.A., Reddi A.H., Moos M.;
RT "Cartilage-derived morphogenetic proteins. New members of the
transforming growth factor-beta superfamily predominantly expressed
in long bones during human embryonic development.";
RL J. Biol. Chem. 269:28227-28234(1994).
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U13661; AAA61416.1; -
HSSP; P18075; IHMP.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb.N.
DR Pfam; PF00019; TGF-beta.1
DR Pfam; PF00688; TGFb-propeptide; 1.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Cytokine; Glycoprotein.
FT NON_TER 1
FT PROPEP <1 316
FT CHAIN 317 436
FT DISULFD 335 401
FT DISULFD 364 433
FT DISULFD 368 435
FT DISULFD 400 400
FT CARBOHYD 27 27
FT CARBOHYD 89 89
SQ SEQUENCE 436 AA; 47873 MW; B0688E12EF8AE91D CRC64;
POTENTIAL.
GROWTH/DIFFERENTIATION FACTOR 6.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
INTERCHAIN (BY SIMILARITY).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
B0688E12EF8AE91D CRC64;

```
Query Match          73.5%; Score 1798; DB 1; Length 436;
Best Local Similarity 80.8%; Pred. No. 4.2e-113;
Matches 354; Conservative 18; Mismatches 46; Indels 18; Gaps 8;

QY 33 SBAELGSGKMSRKGKQAPROSDAGRE--QEPQPRQDDEPR---AQOPRAQEP 86
DB 2 ASAEGLGKMGTRKEGRMPRAPRENATAREPLDRQEPQPRQEPQPRQEPQEP 61

QY 87 GRGPRVPEHYMLSYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDDLSHTPLRQ 146
DB 62 GRGPRVPEHYMLSYRTYSIAEKLGINASFQSSKSANTITSFVDRGLDDLSHTPLRQ 121

QY 147 KYLFVDSMLSDKEELVGAELRFRQAPSAPWGPAGLHVQLVFPCLSPILLDARTLDPOG 206
DB 122 KYLFVDSMLSDKEELVGAELRFRQAPSAPWGPAGLHVQLVFPCLSPILLDARTLDPOG 179

QY 207 APAGWEFVQVQGLRHQPKWQCLLELRAM--GELDAGEAERARQPPQPPDRLSLGF 265
DB 180 APRPGWEFVQVQGLRHQPKWQCLLELRAM--GELDAGEAERARQPPQPPDRLSLGF 239

QY 266 GRVRPQERALLVFTSRQNLFAEMREQLGSA--EAGAGCAGAGGSW-----PPPS 317
DB 240 GRVRPQERALLVFTSRQNLFAEMREQLGSA--EAGAGCAGAGGSW-----PPPS 299

QY 318 GAPDARPNLSPGRRRRRTAFASRHGKRGKSRKRLKPKLHVNFKELGWDDWIIAPLE 377
DB 300 GTPDAGLWSPG-RRRTAFASRHGKRGKSRKRLKPKLHVNFKELGWDDWIIAPLE 358

QY 378 YEAYHCEGVCDFPLRSHLEPHTHAIQTLNMSMDGSPSCCVTKLTPISILYIDAGN 437
DB 359 YEAYHCEGVCDFPLRSHLEPHTHAIQTLNMSMDGSPSCCVTKLTPISILYIDAGN 418

QY 438 NVVYKQYEDMVYSCGR 455
DB 419 NVVYKQYEDMVYSCGR 436

RESULT 2
GDF5_HUMAN
ID GDF5_HUMAN STANDARD; PRT; 501 AA.
AC P43026; Q96SBI;
DT 01-NOV-1995 (Rel. 32, Created)
DE 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Growth/differentiation factor 5 precursor (GDF-5) (Cartilage-derived
DE morphogenetic protein 1) (CDMP-1).
GN GDF5 OR CDMP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95071375; PubMed=7980526;
RA Hoetten G., Neidhardt H., Jacobowsky B., Pohl J.;
RT "Cloning and expression of recombinant human growth/differentiation
RT factor 5.";
RL Biochem. Biophys. Res. Commun. 204:646-652(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular cartilage;
RX MEDLINE=95050604; PubMed=7961761;
RA Chang S., Hoang B., Thomas J.T., Vukicevic S., Luyten F.P.,
RA Ryba N.J.P., Kozak C.A., Reddi A.H., Moos M.;
RT "Cartilage-derived morphogenetic proteins. New members of the
RT transforming growth factor-beta superfamily predominantly expressed
RT in long bones during human embryonic development.";
RL J. Biol. Chem. 269:28227-28234(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
```



```

ID GDF6_MOUSE STANDARD; PRT; 125 AA.
AC P43028;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Growth/differentiation factor 6 precursor (GDF-6) (Fragment).
OS GDF6 OR GDF-6.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=Liver;
RX MEDLINE=941195427; PubMed=8145850;
RA Storm E.E., Huynh T.V., Copeland N.G., Jenkins N.A., Kingsley D.M.,
RA Lee S.-J.;
RT "Limb alterations in brachypodism mice due to mutations in a new
RT member of the TGF beta-superfamily.";
RL Nature 368:639-643(1994).
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08338; AAA18779.1; -.
DR HSSP; P12643; 3BMP.
DR MGD; MGI:95689; Gdf6.
DR InterPro; IPR001839; Tgfb.
DR Pfam; PF00019; TGF-beta; 1.
DR ProDom; PD000357; Tgfb; 1.
DR SMART; SM00204; Tgfb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Cytokine; Glycoprotein.
FT NON_TER 1 1
FT PROPEP <1 5
FT CHAIN 6 125
FT DISULFID 24 90
FT DISULFID 53 122
FT DISULFID 57 124
FT DISULFID 89 89
FT DOMAIN 1 5
FT INTERCHAIN (BY SIMILARITY).
SQ SEQUENCE 125 AA; 14373 MW; 10FA2A5B7748DA32 CRC64;

Query Match 28.2%; Score 691; DB 1; Length 125;
Best Local Similarity 99.2%; Pred. No. 7.7e-40;
Matches 124; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 331 RRRRTAFASRGKRGKSRKSLRCSKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFP 390
DB 1 RRRRTAFASRGKRGKSRKSLRCSKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFP 60
QY 391 LRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPTISILYIDAGNNVYKQYEDMVVE 450
DB 61 LRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPTISILYIDAGNNVYKQYEDMVVE 120
QY 451 SCGR 455
DB 121 SCGR 125

RESULT 5
GDF7_MOUSE
ID GDF7_MOUSE STANDARD; PRT; 151 AA.
AC P43029;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Growth/differentiation factor 7 precursor (GDF-7) (Fragment).
OS GDF7 OR GDF-7.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C; TISSUE=Liver;
RX MEDLINE=941195427; PubMed=8145850;
RA Storm E.E., Huynh T.V., Copeland N.G., Jenkins N.A., Kingsley D.M.,
RA Lee S.-J.;
RT "Limb alterations in brachypodism mice due to mutations in a new
RT member of the TGF beta-superfamily.";
RL Nature 368:639-643(1994).
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U08339; AAA18780.1; -.
DR HSSP; P12643; 3BMP.
DR MGD; MGI:95690; Gdf7.
DR InterPro; IPR001839; Tgfb.
DR Pfam; PF00019; TGF-beta; 1.
DR ProDom; PD000357; Tgfb; 1.
DR SMART; SM00204; Tgfb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Growth factor; Cytokine; Glycoprotein.
FT NON_TER 1 1
FT PROPEP <1 5
FT CHAIN 6 151
FT DISULFID 50 116
FT DISULFID 79 148
FT DISULFID 83 150
FT DISULFID 115 115
FT DOMAIN 1 5
FT POLY-GLY.
FT INTERCHAIN (BY SIMILARITY).
SQ SEQUENCE 151 AA; 15697 MW; 0B496AACB5827759 CRC64;

Query Match 22.4%; Score 549; DB 1; Length 151;
Best Local Similarity 59.5%; Pred. No. 2.8e-30;
Matches 103; Conservative 13; Mismatches 35; Indels 22; Gaps 1;

QY 283 RSQRKLFAMREQLGSAAGPGAGAGSWPPPGAPDARWLPSPGRRRTAFASRH 342
DB 1 RRRRTALAGTGAQSGSGGGGGGGGGGGGGGGG----- 38
QY 343 GKRHGKSRKSLRCSKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAI 402
DB 39 GRGHGRRGRSCSRKSLRCSKPLHVNFKELGWDWIIAPLEYAYHCEGVCDFPLRSHLEPTNHAI 98
QY 403 IQTLNMSMDPGSTPPSCCVPTKLTPTISILYIDAGNNVYKQYEDMVVEGCR 455
DB 99 IQTLNMSMDPGSTPPSCCVPTKLTPTISILYIDAGNNVYKQYEDMVVEGCR 151

RESULT 6
BMPA_XENLA
ID BMPA_XENLA STANDARD; PRT; 398 AA.
AC P25703;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone morphogenetic protein 2-1 precursor (BMP-2-1).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Ocular ciliary epithelium;
RA Wan X.L., Sears J., Chen S., Sears M.;
RT "Cloning and expression of BMP-2/-4 from rabbit ocular ciliary
RT epithelium";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF041421; AAB96785.1; -
DR HSSP; P12643; 3BMP.
DR InterPro; IPR002405; Inhibin_alpha.
DR InterPro; IPR001839; TGFb.
DR InterPro; IPR001111; TGFb_N.
DR Pfam; PF00019; TGF-beta; 1.
DR Pfam; PF00688; TGFb_propeptide; 1.
DR PRINTS; PR00669; INHIBINA.
DR ProDom; PD000357; TGFb; 1.
DR SMART; SM00204; TGFb; 1.
DR PROSITE; PS00250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT SIGNAL 1 23
FT PROPEP 24 281
FT CHAIN 282 395
FT CHAIN 396 408
FT DISULFID 295 360
FT DISULFID 324 392
FT DISULFID 328 394
FT DISULFID 359 359
FT CARBOHYD 134 134
FT CARBOHYD 134 134
FT CARBOHYD 199 199
FT CARBOHYD 337 337
SQ SEQUENCE 395 AA; 44664 MW; 8D1DDCFBAC582496 CRC64;
Query Match 18.4%; Score 449.5; DB 1; Length 395;
Best Local Similarity 29.6%; Pred No. 3.2e-23;
Matches 127; Conservative 68; Mismatches 135; Indels 99; Gaps 19;
QY 58 SDAGREGQEPQPRPODEPRAQQP-----RAQEPGRGPRVVPHEYMLSIYRTYSIAE 109
Db 35 ASSGR----PSPQSDILSEFELRLSMFLGKQRTPSRDAVVP-PYMLDLYRRHS--G 87
QY 110 KLGINA---SFFQSSKSANTISF-VDRGLDLSHTPLR-RQKYLFDVSNLSKDELVCA 164
Db 88 QGAPADPHRLERASRANTVRSFHEESLELPETSGKTRFRFNLTSIPPEFITS 147
QY 165 ELRLFRQAPSAPWGPAGPLH-VQLFPCLSP-----LLDARTLDPQAGPAGWE 213
Db 148 ELQVFEQMEQALGDDSGFHRINIVYIIPATANSKFPATRLDRLVYNQ-----TSRWE 204
QY 214 VFDVWGLRHQPKQLCLERANGELDAGEAEARAGPQQPPPPDLRLSLGFRVRVP-- 271
Db 205 SFDVTPAVNR--W-----TAQHANGFV-VEVTHLEERQGVSKRHRVIRLSLHPDE 253
QY 272 ---PQERALLVFTSRQKNLFAEMREQLSAFAAGPGAGAGSGSWPPPSGADPARPWLPS 328
Db 254 HWSQIRPLLVTF-----GHDGKGHLH----- 276
QY 329 PGRRRRTAFASRHKRHKKSRLR--CSKKPLHVNFKELGDDWDTIAPLEYEAVHCEGV 386
Db 277 --RRKROA-----KHQRKRLKSSCRHPLYVDFSDVGNWDNIVAPPGTHAFYCHGE 327

QY 387 CDFPLRSHLEPTNHAIQTLMNSMDPGSTPPSCVPTKLTPTISILYIDAGNNVYKQYED 446
Db 328 CDFPLADHLNSTNHAIQTLMNSVN-SKIPKACVPTELTSALSMYLDENKVKVLYNQD 386
QY 447 MVVESCGR 455
Db 387 MVVEGCGCR 395
RESULT 10
BMP4_HUMAN
ID BMP4_HUMAN STANDARD; PRT; 408 AA.
AC P12644; Q9UM80;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2B).
GN BMP4 OR BMP2B OR DVRA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89072730; PubMed=3201241;
RA Wozney J.M., Rosen V., Celeste A.J., Mitsock L.M., Whitters M.J.,
RA Kriz R.W., Hewick R.M., Wang E.A.;
RT "Novel regulators of bone formation: molecular clones and
RT activities";
RL Science 242:1528-1534(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98368971; PubMed=9701626;
RA Shore E.M., Xu M., Shah P.B., Janoff H.B., Hahn G.V., Deardorff M.A.,
RA Sovinsky L., Spinner N.B., Zasloff M.A., Wozney J.M., Kaplan F.S.;
RT "The human bone morphogenetic protein 4 (BMP-4) gene: molecular
RT structure and transcriptional regulation.";
RL Calcif. Tissue Int. 63:221-229(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Durdenum;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 7-408 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96063017; PubMed=7579580;
RA Oida S., Iimura T., Maruoka Y., Takeda K., Sasaki S.;
RT "Cloning and sequence of bone morphogenetic protein 4 (BMP-4) from a
RT human placental cDNA library";
RL DNA Seq. 5:273-275(1995).
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION ALSO ACT IN
CC MESODERM INDUCTION, TOOTH DEVELOPMENT, LIMB FORMATION AND FRACTURE
CC REPAIR.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE LUNGS AND LOWER LEVELS SEEN
CC IN THE KIDNEY. PRESENT ALSO IN NORMAL AND NEOPLASTIC PROSTATE
CC TISSUES, AND PROSTATE CANCER CELL LINES.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M22490; AAA51835.1; -
DR EMBL; U43842; AAC72278.1; -
DR EMBL; BC020546; AAB0546.1; -
DR EMBL; D30751; BAA06410.1; -

```
DR PIR: C37278; C37278.
DR HSP: P12643; 3BMP.
DR Genew: HGNC:1071; BMP4.
DR MIM: 112262; -.
DR InterPro: IPR001839; TGFb.
DR InterPro: IPR001111; TGFb_N.
DR Pfam: PF00019; TGF-beta; 1.
DR Pfam: PF00688; TGFb-propeptide; 1.
DR ProDom: PD000357; TGFb; 1.
DR SMART: SM00204; TGFb; 1.
DR PROSITE: PS00250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT SIGNAL 1 19
FT PROPEP 20 292
FT CHAIN 293 408
FT DISULFID 308 373
FT DISULFID 337 405
FT DISULFID 341 407
FT DISULFID 372 372
FT CARBOHYD 143 143
FT CARBOHYD 208 208
FT CARBOHYD 350 350
FT CARBOHYD 365 365
FT CONFLICT 152 152
SQ SEQUENCE 408 AA; 46555 MW; 79B01179DBB98204 CRC64;

Query Match 18.0%; Score 441.5; DB 1; Length 408;
Best Local Similarity 28.0%; Pred. No. 1.1e-22;
Matches 135; Conservative 70; Mismatches 164; Indels 113; Gaps 19;

QY 7 LLSAVLFLSLDLRGFQOASISSCSAELGTGK-MRSREGKMKQAPRDSDA--GR 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
7 MLMVLLCQVL--LGGASHASLIPETGKKVAEIQGHAGRGSGOSHELLRFEATLQM 64
QY 63 EQGEQPPQDPEPRAQPPQRAQPPGPRVVPHEMYLSIYRYSIAEK-----LGINAS 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
65 FGLRRPQPSKS-----AVIP-DYMRDLYRLQSGEEDEEIQHSTGLEYP 107
QY 117 PFOSSKSANTITSF-VDRGLDLSHTPLRRQ-KYLFVDSMLSKDELGAELRLRQAPS 174
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
108 ERPASR-ANTVRSFHEEHELENIPGTSNSAFRFLNLSIPENEVISAEALRLFRE--Q 164
QY 175 APWGP--PAGPLVHVLFPCLSP-----LLDARTLDPOGAPPAGNVEFVDMOGL 221
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
165 VDQGPDPWGEGRFINIYEMKPPAEVVPGLHTRLLDTRLVHVN---VTRWETEDV--- 217
QY 222 RHQPKQLCLELRAWGLDAGEAEARGPQPPDLRLSLGFGRRVPPPOERALLVVF 281
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
218 -----SPAVLR---WTRKQPNYGLAIEVTH 240
QY 282 TRSQKNLFAEMREQLGSAEAGAGAGSGWPPPSGAPDARPLPSPGR-----RR 333
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
241 LHQTRTHOGHVRSLSLPQGS-----GNW-----AQLRPLLYTGHDRGHALTRR 287
QY 334 RTAFASRHKRGKSRKSLRCSKPLHVNFKELGWDWDWIAPLEYAYHCEGCVDFPLRS 393
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
288 RRAKSPKHQSQRARKKNCRRHSLSYVDFSDVGNWDWIVAPPGVQAFYCHGDCPEPLAD 347
QY 394 HLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISLYIDAGNVVYQYEDMVVESC 453
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
348 HLNSTNHAIQTVLNSVN-SSIPKACCVPTSLSMYLDYDKVWLKLVQEMVVEGCG 406
QY 454 CR 455
Db : :
407 CR 408

RESULT 11
BMP4_MOUSE
ID BMP4_MOUSE
AC P21275;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
```

```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2B).
GN BMP4 OR BMP-4 OR DVR-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Dickinson M.E., Van der Meer-De Jong R., Hogan B.L.M.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93282803; PubMed=8507180;
RA Kurihara T., Kitamura K., Takao K., Nakazato H.;
RT "Murine bone morphogenetic protein-4 gene: existence of multiple
promoters and exons for the 5'-untranslated region.";
RL Biochem. Biophys. Res. Commun. 192:1049-1056(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93365172; PubMed=8358941;
RA Takao K., Yoshikawa H., Hashimoto J., Masuhara K., Miyamoto S.,
RA Suzuki S., Ono K., Matsui M., Oikawa S., Tsuruoka N.;
RT "Gene cloning and expression of a bone morphogenetic protein derived
from a murine osteosarcoma.";
RL Clin. Orthop. Relat. Res. 294:344-352(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Liver;
RX MEDLINE=96081880; PubMed=7499338;
RA Feng J.Q., Chen D., Cooney A.J., Tsai M., Harris M.A., Tsai S.Y.,
RA Feng M., Mundy G.R., Harris S.E.;
RT "The mouse bone morphogenetic protein-4 gene. Analysis of promoter
utilization in fetal rat calvarial osteoblasts and regulation by
COUP-TF1 orphan receptor.";
RL J. Biol. Chem. 270:28364-28373(1995).
RN [5]
RP SEQUENCE OF 241-408 FROM N.A.
RX MEDLINE=90228966; PubMed=1970330;
RA Dickinson M.E., Kobrin M.S., Silan C.M., Kingsley D.M., Justice M.J.,
RA Miller D.A., Cecil J.D., Lock L.F., Lee A., Buchberg A.M.,
RA Siracusa L.D., Lyons K.M., Derynck R., Hogan B.L.M., Copeland N.G.,
RA Jenkins N.A.;
RT "Chromosomal localization of seven members of the murine TGF-beta
superfamily suggests close linkage to several morphogenetic mutant
loci";
RL Genomics 6:505-520(1990).
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56848; CAA40179.1; -.
DR EMBL: S65032; AAB28021.1; -.
DR EMBL: L47480; AAC37698.1; ALT_INIT.
DR EMBL: D14814; BAA03555.1; -.
DR PIR: B34201; B34201.
DR PIR: S29523; S29523.
DR HSSP: P12643; 3BMP.
DR MGD: MGI:88180; Bmp4.
DR InterPro: IPR001839; TGFb.
DR InterPro: IPR001111; TGFb_N.
DR Pfam: PF00019; TGF-beta; 1.
DR Pfam: PF00688; TGFb-propeptide; 1.
DR ProDom: PD000357; TGFb; 1.
```



```
BMP4_RAT
ID BMP4_RAT          STANDARD;          PRT;          408 AA.
AC Q06826;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bone morphogenetic protein 4 precursor (BMP-4) (BMP-2b).
GN BMP4 OR BMP-4 OR DVR-4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93385158; PubMed=8373807;
RA Chen D., Feng J.Q., Feng M., Harris M.A., Mundy G.R., Harris S.E.;
RT "Cloning and sequence of bone morphogenetic protein 4 cDNA from fetal
RT rat calvarial cell.";
RL Biochim. Biophys. Acta 1174:289-292(1993).
CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 222607; CAA80329.1; .
DR PIR: S33173; S33173.
DR PIR: S38343; S38343.
DR HSP: P12643; 3BPB.
DR InterPro: IPR001839; TGFb.
DR InterPro: IPR001111; TGFb.N.
DR Pfam: PF00019; TGF-beta; 1.
DR Pfam: PF00688; TGFb.propeptide; 1.
DR ProDom: PD000357; TGFb; 1.
DR SMART: SM00204; TGFb; 1.
DR PROSITE: PS00250; TGF_BETA_1; 1.
KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
FT SIGNAL 1 19
FT PROPEP 20 292
FT CHAIN 293 408
FT DISULFID 308 373
FT DISULFID 337 405
FT DISULFID 341 407
FT DISULFID 372 372
FT CARBOHYD 144 144
FT CARBOHYD 209 209
FT CARBOHYD 350 350
FT CARBOHYD 365 365
SQ SEQUENCE 408 AA; 46540 MW; 61E92B4B8D5624F3 CRC64;

Query Match 17.7%; Score 432.5; DB 1; Length 408;
Best Local Similarity 29.1%; Pred. No. 4.5e-22;
Matches 138; Conservative 62; Mismatches 168; Indels 107; Gaps 19;

QY 20 LPGFOAQSTSSCSAEALGSKGMRSEKGMORAPRDSAG--REGQEPQPRQDEPRA 77
DB 2 IFGNRLAVLVLCVQLGGATDASLPETGKKVAIEIQHAGGRSGSHELLRFEATL 61
QY 78 QQP-----RAQEPGPRVVPVHEMYLSIYTYTSIABK-----IGINASFQSSKS-ANTIT 128
DB 62 LQMFGLRRRPPSKS-AVIP-DYMRDLRLQSGEEEBEQSQGTGLEYPERPASRANTVR 119
QY 129 SFVDRGLDLSHTPLRRQ-----KYLFDVYSMLSDKEELVGAELRLFR----QAPSAPWG-- 178
DB 120 SPFHE--EHLENIPGTSSESAFRFFFNLSIPENVEISSAELRLRFRQVDGPDWEOGFH 177
```

```
QY 179 -----PPAG---PLHVQLFPCLSPLLLDARTLDPOGAPPAGWEVFDVWQGLRHQPWK 227
DB 178 RNIYEVKPPAEMVPGH-----LITRLDTRLVRHN---VTRWETFDV----- 218
QY 228 QLCLELRAAWGELDAGEAFARAGPQQPPPDLSLIGFGRVRRPQPQERALLVVFTSRQK 287
DB 219 -----SPAVLR---WTREKQPNYGLAIEVTHLHQTRT 247
QY 288 NLFAMREOLGSAEAGPGAGAGSGSWPPSPGAPDARPLWPSGR-----RRRTAFAS 340
DB 248 HQGHVTRISRSLPQGS-----GNW-----AQLRPLLVTFEGHGRHTLRRRAKSP 294
QY 341 RHGKRHGKSRCLKPLHVNFKELGWDDWTIAPLEYEAHCEGVCDPFLSHLEPTNH 400
DB 295 KHPORSRKKNKNCRRHSLYVDFSDVGNWDWIAPPGYQAFYCHGDCPPPLADLNSTNH 354
QY 401 AIIQTLMSMDPGSTPCCVPTKLTPIISILYIDAGNNVNYKYQYEDVMVYSGCGR 455
DB 355 AIVQTLVNSVN-SSIPKACCVPTELSALSMYLDVYKVLKNYQEMVYEGGCR 408

RESULT 14
BMP2_HUMAN
ID BMP2_HUMAN          STANDARD;          PRT;          396 AA.
AC P12643;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bone morphogenetic protein 2 precursor (BMP-2) (BMP-2a).
GN BMP2 OR BMP2A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89072730; PubMed=3201241;
RA Wozney J.M., Rosen V., Celeste A.J., Mitscock L.M., Whitters M.J.,
RA Kriz R.W., Hewick R.M., Wang E.A.;
RT "Novel regulators of bone formation: molecular clones and
RT activities.";
RL Science 242:1528-1534(1988).
RN [2]
RP SEQUENCE FROM N.A.
RA Shore E.M., Xu M.-Q., Calvert G., Moriatias J., Kaplan F.S.;
RT "Human bone morphogenetic protein 2 (BMP-2) genomic DNA sequence.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Ciamp M., Clark A.P., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leivasalio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McIlroy K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Suliscon J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
```

RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 292-396.
 RX MEDLINE:99175323; PubMed:10074410;
 RA Scheufler C., Sebald W., Huelsmeyer M.;
 RT "Crystal structure of human bone morphogenetic protein-2 at 2.7 A
 RT resolution.";
 RL J. Mol. Biol. 287:103-115(1999).
 CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION.
 CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -!- TISSUE SPECIFICITY: PARTICULARLY ABUNDANT IN LUNG, SPLEEN AND
 CC COLON AND IN LOW BUT SIGNIFICANT LEVELS IN HEART, BRAIN, PLACENTA,
 CC LIVER, SKELETAL MUSCLE, KIDNEY, PANCREAS, PROSTATE, OVARY AND
 CC SMALL INTESTINE.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M22489; AAA51834.1; .
 DR EMBL; AF040249; AAF21646.1; .
 DR EMBL; AL035668; CAB82007.1; .
 DR PIR; B37278; B37278.
 DR PDB; 3BMP; 12-MAR-00.
 DR Genew; HGNC:1069; BMP2.
 DR MIM; 112261; .
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb.N.
 DR Pfam; PF000019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb-propeptide; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein;
 KW 3D-structure. 1 23 POTENTIAL.
 FT SIGNAL 1 23
 FT PROPEP 24 282
 FT CHAIN 283 396
 FT DISULFID 296 361
 FT DISULFID 325 393
 FT DISULFID 329 395
 FT DISULFID 360 360
 FT CARBOHYD 135 135
 FT CARBOHYD 163 163
 FT CARBOHYD 164 164
 FT CARBOHYD 200 200
 FT SEQUENCE 396 AA; 44702 MW; 20653A3987B25E60 CRC64;
 Query Match 17.6%; Score 431.5; DB 1; Length 396;
 Best Local Similarity 29.2%; Pred. No. 5.1e-22;
 Matches 130; Conservative 67; Mismatches 139; Indels 109; Gaps 20;
 QY 49 GKWRAPRDSADGREGEPQPRQDEPRAQPP-----RAQEPGGRPRVVPHEVMLS 100
 Db 23 GLVPELGRKKFAAASGRPSQSDSEFELRLLSMFGKLRQTPPSRDVVP-PYMLD 81
 QY 101 IYRTSYAEKLGINAS-----PFQSKSANTITSF-VDRGLDLSHTPLR-RQKYLFDV 152
 Db 82 LYRHS-----GQPGPAPDHLRLERASRANTVRSFHEESLEELPESKTTTRFFNL 136
 QY 153 SMLSDKEELYAEALRFLQAPSAPGPPAGPLH-VOLFCLSP-----LILDAT 201
 Db 137 SSIPTEEFITSAELQVREQMDALGNSSFFHRIYELIKPATANSKFPVRLDTRL 196
 QY 202 LDPQAGPAGWEVDVWQGLRHQFWKQLCLELRAANGELDAGEAEARARGPQPPPDRL 261

Db 197 VN-QNA--SRWESFDVTPAYMR--W-----TAQGHANHGFFVVEVAHLEEKQ----- 237
 QY 262 SLGFGRRVRPQERALLVWTRSORKNLFARERQOLGSAEAGPGACGAGSGWPPSPGAPD 321
 Db 238 --GVSKR-----HVRISRSLHQD-----EHSW-----SQ 259
 QY 322 ARWLPSPG-----RRRRRTAFASRHGKRHKSRRLR--CSKKPLHVNFKELQWDD 370
 Db 260 IRPLLVTFGHDGKGPLHKREKQA-----XHKQKRLKSSCKRHPLYVDFSDVGWDD 312
 QY 371 WITAPLEYAYHCEGVCDFPLRSHLEPTNHAITLNSMDPGSTPPSCCVPKLTISI 430
 Db 313 WIWAPPGYHAFYHCEGCPFLADHLNTHAIVQTLNSVN-SKIPKACCCVPTLSAISM 371
 QY 431 LYIDAGNNVYKQVEDMVVESCGR 455
 Db 372 LYLDENEKVKLVQYDMVVEGCGCR 396
 RESULT 15
 BMP4_RABIT
 ID BMP4_RABIT STANDARD; PRT; 409 AA.
 AC O46576;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bone morphogenetic protein 4 precursor (BMP-4).
 GN BMP4 OR BMP-4.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 ON NCBI_TaxID=9986;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=New Zealand white; TISSUE=Ocular ciliary epithelium;
 RA Wan X.L., Sears J., Chen S., Sears M.;
 RT "Cloning and expression of BMP-2/-4 from rabbit ocular ciliary
 RT epithelium.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION (BY SIMILARITY).
 CC -!- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: SECRETED INTO THE EXTRACELLULAR MATRIX.
 CC -!- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF042497; AAB97467.1; .
 DR HSSP; P12643; 3BMP.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb.N.
 DR Pfam; PF000019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb-propeptide; 1.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Signal; Growth factor; Cytokine; Bone; Cartilage; Glycoprotein.
 FT SIGNAL 1 19 POTENTIAL.
 FT PROPEP 20 293
 FT CHAIN 294 409
 FT DISULFID 309 374
 FT DISULFID 338 406
 FT DISULFID 342 408
 FT DISULFID 373 373
 FT CARBOHYD 144 144
 FT CARBOHYD 209 209
 FT CARBOHYD 351 351
 BY SIMILARITY.
 BONE MORPHOGENETIC PROTEIN 4.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 INTERCHAIN (BY SIMILARITY).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 366 366 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 409 AA; 46641 MW; 35557561447AD625 CRC64;

Query Match      17.6%; Score 430; DB 1; Length 409;
Best Local Similarity 27.7%; Pred. No. 6.7e-22;
Matches 134; Conservative 68; Mismatches 167; Indels 114; Gaps 18;

QY 7 LLSAVFLISFLMDLPGFQOASISSSCSSAELGSTKG-MRSRKREGKMQRAPRDSDA---GR 62
Db 7 LLMVLLCQVL--LGASHASLIPETGKKVAIEIQHAGRRSGQSGSHELLRFEATLLQM 64

QY 63 EGQEPQRFQDEPRAQQPRAQPPGRVPRVHEYMLSIYRTYSIAEKL-----GINA 115
Db 65 FGLRRHPQPSKS-----AVIP-DYMRDLYRLQSGEEEEEQMPSGGLE 107

QY 116 SFFQSKSANTITSF-VDRGLDDLSTPLRRQ-KYLFVSMLSKDEELVGAELRLFRQAP 173
Db 108 PERPASR-ANTVRSFHHEHLENIPGTSSENSAFRLNLSIPENEAISSAELRLFRE-- 164

QY 174 SAPWGP--PAGPLHVOLFCLSP-----LLLDARTLDPOCAPPAGWEFVDVWOG 220
Db 165 QVDQGDWNERGFHRINIYEVMPPPAEAVPGHLITRLDTRLVHHN---VTRWETFDV--- 218

QY 221 LRHQPWKQLCLELRAAWGELDAGEAEARARGPQQPPPPDLRLSLGFGRRVRPPQOERALLVY 280
Db 219 -----SPAVLR---WTREKQPNHGLAVEVT 240

QY 281 FTRSQKNLFAEMREQLGSAEAGPCAGAGSWPPPSGAPDARPWLPSGR-----R 332
Db 241 HFHHTTHQGQHVRLSRSLQSGDWA-----QFRPLLVTFGHDGRGHALT 287

QY 333 RRTAFASRHGKRHKSRKRLSCSKPLHYNFKELGWDWTIAPLEYEAYHCEGVCDFPLR 392
Db 288 RRAKSLAKHHQPQARKKKNCRHALYVDFSDVGNWDNIIVAPPGYQAFYCHGDCFPPLA 347

QY 393 SHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILIYIDAGNNVYKQYEDMYVESC 452
Db 348 DHFNSTNHAIVQTLVNSVN-SSIPKACCVPTELSAISMLYLDDEYDKVLKNYQEMVVEGC 406

QY 453 GCR 455
Db 407 GCR 409
```

Search completed: November 25, 2002, 02:56:09
Job time : 38 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 00:20:15 : Search time 3689 Seconds
(without alignments)
11241.935 Million cell updates/sec

Title: US-09-825-751A-19
Perfect score: 1425
Sequence: 1 cctccgggagacgcagcca.....tttccgcgcgttgcccg 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pi.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rod.*
- 36: em_hgt_mam.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgtgo_hum.*
- 40: em_hgtgo_mus.*
- 41: em_hgtgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1425	100.0	1425	6	AX277361	Sequence
2	1425	100.0	1425	6	AX465491	Sequence
3	1421.8	99.8	2791	6	AX351025	Sequence
C 4	991	69.5	156873	9	AC015998	Homo sapi
C 5	991	69.5	230879	9	AP003465	Homo sapi
6	967	69.3	1171	6	I62379	Sequence 25
7	935	65.6	184987	2	AC126226	Papio cyn
8	928.2	65.1	1308	4	BTU13661	Bos taurus
9	928.2	65.1	1308	6	AX277426	Sequence
10	885.8	62.2	143312	2	AC126238	Felis cat
C 11	798.8	56.1	219978	2	AC126232	Bos tauru
C 12	765.8	53.7	209733	10	AC058786	Mus muscu
13	755.4	53.0	326540	2	AL732476	Mus muscu
14	749.8	52.6	167312	2	AC116927	Rattus no
15	581.8	40.8	1046	6	I62382	Sequence 31
C 16	426.8	30.0	253694	2	AC084005	Homo sapi
C 17	389.8	27.4	177269	2	AC114354	Rattus no
C 18	377.6	26.5	136505	2	AC126926	Felis cat
C 19	360.2	25.3	1880	5	AF155125	Xenopus l
C 20	343.4	24.1	138738	2	AC126234	Canis fam
C 21	343	24.1	179366	2	AC114503	Rattus no
22	320.6	22.5	375	10	MMU08338	Mus musculu
C 23	296	20.8	134623	2	AC126235	Canis fam
24	285.8	20.1	1354	5	DRDYNAMO	X99769 D.rerio mrn
C 25	270.6	19.0	381	5	AF089086	Gallus ga
C 26	267	18.7	177514	2	AC126919	Bos tauru
27	262.4	18.4	2169	9	AF254568S2	Cercopith
28	259.2	18.2	1345	6	I62383	Sequence 33
29	259.2	18.2	206021	9	AC012065	Homo sapi
30	255	17.9	183402	2	AC126231	Bos tauru
31	253.6	17.8	2694	9	AF254567	Macaca fa
32	253	17.8	2323	9	HSU13660	Human carti
33	253	17.8	2703	6	AX3194	Sequence 1
34	253	17.8	2703	6	A59669	Sequence 2
35	253	17.8	2703	6	A60247	Sequence 1
36	253	17.8	2703	6	A79612	Sequence 2
37	253	17.8	2703	6	AR091289	Sequence
38	253	17.8	2703	6	E12603	CDNA encodi
39	253	17.8	2703	6	E14291	DNA encodin
40	253	17.8	2703	6	E15606	CDNA encodi
41	253	17.8	2703	6	HS606	CDNA encodi
C 42	253	17.8	151903	9	HSJ47704	Human DNA
43	251.6	17.7	1207	6	AR137664	Sequence
44	251.6	17.7	1207	6	I62365	Sequence 3
45	251.4	17.6	387	6	A79614	Sequence 4

ALIGNMENTS

RESULT 1	AX277361	Sequence 19	1425 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	AX277361	from Patent	WO0174897.			
DEFINITION	AX277361					
ACCESSION	AX277361					
VERSION	AX277361.1					
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Vernet,C.A., Burgess,C.E., Fernandes,E., Taupier,R.J., Quinn,K.E., Spytek,K.A., Rastelli,L. and Herrmann,J.L.					
TITLE	Novel proteins and nucleic acids encoding same					

CDS	/organism="Homo sapiens" /db_xref="taxon:9606" 31..1398 /note="unnamed protein product" /codon_start=1 /protein_id="CAD42373.1" /db_xref="GI:21899854" /translation="MDTRVLLSAVFLISFLWDLPGFQQAISISSCSAELGSKGMR SRKEGMRAPRDSDAQREGQPPQDEPRAQPPQPPRQVPRVPHYMLSIY RTYSIAELGINSFQSKSANTITFVDRGLDDLSTPLRQKYLFDVSMLSKDEE LVGLRFPQKQAPAGPAGPLHVOLFPLCLSPDLLDARTLDPOGAPAGWVEVDW QGLRHQPKQLCLERAAWGLDAEAEARAGPQPPDLRLSLGFRVRRPQERA LLVFTRSQRNLFAEMREQLGSAEAGPGAGAGBSWPPSPGADARWPLSPGRRR RTAFASRHGRKGRSLRCKPLHVNFKELGDDWIIAPLEYEAYHCEGVDFPLR SLEPTNHALIOTLMSMDPGSTPPSCVPTKLTPISILYIDAGNNVYKQYEDMVE SCGR"		234 a	488 c	469 g	234 t
BASE COUNT						
ORIGIN						
	Query Match 100.0%; Score 1425; DB 6; Length 1425; Best Local Similarity 100.0%; Pred. No. 3.9e-215; Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	CTCCTGGGAGAGCAGCCACTTGCCTCCCATGGATACCTCCAGGGTCTGCTCTCGGCC	60			
Db	1	CTCCTGGGAGAGCAGCCACTTGCCTCCCATGGATACCTCCAGGGTCTGCTCTCGGCC	60			
QY	61	GTCTTCTCATCAGTTTCTGTGGGATTTCGCCGTTTCCAGCAGGCTTCCATCTCATCC	120			
Db	61	GTCTTCTCATCAGTTTCTGTGGGATTTCGCCGTTTCCAGCAGGCTTCCATCTCATCC	120			
QY	121	TGCTGTTCTGCCGAGCTGGGTTCCACCAGGCGCATGCGAAGCCGCAAGGAAGCAAG	180			
Db	121	TGCTGTTCTGCCGAGCTGGGTTCCACCAGGCGCATGCGAAGCCGCAAGGAAGCAAG	180			
QY	181	ATGCACGGCGCGCGGCACAGTACGCGGCGCGGAGGCGCCAGGACACACAGCGCGG	240			
Db	181	ATGCACGGCGCGCGGCACAGTACGCGGCGCGGAGGCGCCAGGAAACACACGCGCGG	240			
QY	241	CCTCAGGACGAACCCCGGCTCAGCAGGCCCGCGGCGCAGGAGCCGCGAGGGGTCG	300			
Db	241	CCTCAGGACGAACCCCGGCTCAGCAGGCCCGCGGCGCAGGAGCCGCGAGGGGTCG	300			
QY	301	CGCGTGGTCCCGACAGTACATGTGTCAATCTACAGGACTTACTTCATCGCTGAGAAG	360			
Db	301	CGCGTGGTCCCGACAGTACATGTGTCAATCTACAGGACTTACTTCATCGCTGAGAAG	360			
QY	361	CTGGGATCAATGCGCAGCTTTTTCAGTCTTCCAAGTCGGCTAATACCATCAGCAGTTT	420			
Db	361	CTGGGATCAATGCGCAGCTTTTTCAGTCTTCCAAGTCGGCTAATACCATCAGCAGTTT	420			
QY	421	GTAGACAGGGGACTAGACGATCTCTCGACACTCTCTCTCCGGAGACAGAATTTTGT	480			
Db	421	GTAGACAGGGGACTAGACGATCTCTCGACACTCTCTCTCCGGAGACAGAATTTTGT	480			
QY	481	GATGTGTCATGCTCTCAGACAAAGAAGAGCTGTGTGGCGCGGAGCTCGGCTCTTTCGC	540			
Db	481	GATGTGTCATGCTCTCAGACAAAGAAGAGCTGTGTGGCGCGGAGCTCGGCTCTTTCGC	540			
QY	541	CAGGCGCCCTCAGCGCCTTGGGGCCACACAGCGGGCGCTCCAGTGCAGCTCTTCCT	600			
Db	541	CAGGCGCCCTCAGCGCCTTGGGGCCACACAGCGGGCGCTCCAGTGCAGCTCTTCCT	600			
QY	601	TGCTTTCCGCCCTACTGCTGACCGCGGACCCCTGAGCCCGCAGGGGGCGCCCGCGCC	660			
Db	601	TGCTTTCCGCCCTACTGCTGACCGCGGACCCCTGAGCCCGCAGGGGGCGCCCGCGCC	660			
QY	661	GCGTGGGAAGTCTTCAGCTGTGGCAGGGCCTTGCGCCACACAGCCCTGGAAGCAGCTGTC	720			
Db	661	GCGTGGGAAGTCTTCAGCTGTGGCAGGGCCTTGCGCCACACAGCCCTGGAAGCAGCTGTC	720			
QY	721	TTGGAGCTCGCGGCGCATGTGGGCGAGCTGACGCGCGGGAGCGCGAGCGCGCGCGG	780			

Query Match		99.8%; Score 1421.8; DB 6; Length 2791;
Best Local Similarity		99.9%; Pred. No. 1.1e-214;
Matches 1423; Conservative		0; Mismatches 2; Indels 0; Gaps 0;
QY	1	CTCCTGGGAGACGACGACACTTGGCCCGCATGGATACCTCCAGAGGTCTGCTCTCGGCG 60
Db	999	CTCCTGGGAGACGACGACACTTGGCCCGCATGGATACCTCCAGAGGTCTGCTCTCGGCG 1058
QY	61	GTCTTCTCATAGTTTCTGTGGATTTCGCCGTTTCACAGAGCTTCCATCTCATCC 120
Db	1059	GTCTTCTCATAGTTTCTGTGGATTTCGCCGTTTCACAGAGCTTCCATCTCATCC 1118
QY	121	TCCTGTTCGTCGCGAGCTGGGTTCACCAAGGCGATCGAAGCGCGCAAGGAAGCAAG 180
Db	1119	TCCTGTTCGTCGCGAGCTGGGTTCACCAAGGCGATCGAAGCGCGCAAGGAAGCAAG 1178
QY	181	ATGCAGCGGGCGCGCGCGACAGTACGCGGGCGGGAGGGCCAGGAACACAGCGCGGG 240
Db	1179	ATGCAGCGGGCGCGCGCGACAGTACGCGGGCGGGAGGGCCAGGAACACAGCGCGGG 1238
QY	241	CCTCAGGACGAACCCCGGCTCAGCAGCCCGGGCGCAGGAGCCGCCAGGAGGGTCGG 300
Db	1239	CCTCAGGACGAACCCCGGCTCAGCAGCCCGGGCGCAGGAGCCGCCAGGAGGGTCGG 1298
QY	301	CGCGTGGTGCCTCCACGAGTACATGCTGTCAATCTACAGGACTTACTCCATCGTGAGAAG 360
Db	1299	CGCGTGGTGCCTCCACGAGTACATGCTGTCAATCTACAGGACTTACTCCATCGTGAGAAG 1358
QY	361	CTGGGATCAATGCCAGCTTTTTCAGTCTTCCAGTTCGGCTAATACGATCACCAGCTTT 420
Db	1359	CTGGGATCAATGCCAGCTTTTTCAGTCTTCCAGTTCGGCTAATACGATCACCAGCTTT 1418
QY	421	GTAGCAGGGGACTAGACCATCTCTCGCACACTCCTCTCCGGAGACAGAGTATTGTGTT 480
Db	1419	GTAGCAGGGGACTAGACCATCTCTCGCACACTCCTCTCCGGAGACAGAGTATTGTGTT 1478
QY	481	GATGTGTCATGCTCTCAGACAAAGAGAGCTGGTGGCGCGGAGCTGGGCTCTTTTCGC 540
Db	1479	GATGTGTCATGCTCTCAGACAAAGAGAGCTGGTGGCGCGGAGCTGGGCTCTTTTCGC 1538
QY	541	CAGGCGCCTCAGCGCCTTGGGGCCACAGCGGGCGGCTCAGGTGAGCTCTTCCCT 600
Db	1539	CAGGCGCCTCAGCGCCTTGGGGCCACAGCGGGCGGCTCAGGTGAGCTCTTCCCT 1598
QY	601	TGCCTTTCGCCCTACTGCTGGACGCGCGGACCTTGGACCCCGAGGGCGCGCGGCG 660
Db	1599	TGCCTTTCGCCCTACTGCTGGACGCGCGGACCTTGGACCCCGAGGGCGCGCGGCGG 1658
QY	661	GGCTGGGAAGTCTTGACGTGTGGCAGGGCCCTGCGCCACACAGCCCTTGAAGCAGCTGTGC 720
Db	1659	GGCTGGGAAGTCTTGACGTGTGGCAGGGCCCTGCGCCACACAGCCCTTGAAGCAGCTGTGC 1718
QY	721	TTGGAGCTCGGGCGCATATGGGGCAGCTGGACGCGGGGAGGCGCAGCGCGCGGG 780
Db	1719	TTGGAGCTCGGGCGCATATGGGGCAGCTGGACGCGGGGAGGCGCAGCGCGCGGG 1778
QY	781	GGACCCCGACCAACCGCGCCCGGACCTGCGAGTCTGGGCTTGGCGGAGGCTGGG 840
Db	1779	GGACCCCGACCAACCGCGCCCGGACCTGCGAGTCTGGGCTTGGCGGAGGCTGGG 1838
QY	841	CCTCCCCAGGAGCGGGCCCTGCTGGTGTGTTATTCACAGATCCACGCAAGAACCTGTTC 900
Db	1839	CCTCCCCAGGAGCGGGCCCTGCTGGTGTGTTATTCACAGATCCACGCAAGAACCTGTTC 1898
QY	901	GCAGAGATCGCGAGCAGCTGGGCTGGCCGAGGCTGGGGGCCCGGGCGCGGCGCCGAG 960
Db	1899	GCAGAGATCGCGAGCAGCTGGGCTGGCCGAGGCTGGGGGCCCGGGCGCGGCGCCGAG 1958
QY	961	GGGTCTGTGGCCCGCGCTGGCGGCGCCCGGATGCCAGGCTTGGCTGCGCTCGCCCGCGC 1020
Db	1959	GGGTCTGTGGCCCGCGCTGGCGGCGCCCGGATGCCAGGCTTGGCTGCGCTCGCCCGCGC 2018
QY	1021	CGCCGGCGGCGCGCACGCGCCTTCGCCAGTGCCTCATGGCAGCGGCACGGCAAGAAGTCC 1080

Db	2019	CGCGGCGCGGCGCACGCGCTTCGCCAGTGCCTAGCGACGGCGACGGCAGAGTCC 2078
QY	1081	AGGCTACGCTGAGCAAGAAGCCCTGACGTGAACCTTCAAGAGAGTGGGCTGGGACGAC 1140
Db	2079	AGGCTACGCTGAGCAAGAAGCCCTGACGTGAACCTTCAAGAGAGTGGGCTGGGACGAC 2138
QY	1141	TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCCG 1200
Db	2139	TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCCG 2198
QY	1201	CTGCGCTCGACCTGAGCGCCCAACACCGCATCATCAGACGCTGATGAATCCATG 1260
Db	2199	CTGCGCTCGACCTGAGCGCCCAACACCGCATCATCAGACGCTGATGAATCCATG 2258
QY	1261	GACCCGGGCTCACCCCGCCAGCTGCTGCGTGGCCCAACAAATTGACTCCCATCAGCAT 1320
Db	2259	GACCCGGGCTCACCCCGCCAGCTGCTGCGTGGCCCAACAAATTGACTCCCATCAGCAT 2318
QY	1321	CTATACATCGACGCGGGCAATAATGTGTTCTACAAGCATGAGGACATGGTGGTGAG 1380
Db	2319	CTATACATCGACGCGGGCAATAATGTGTTCTACAAGCATGAGGACATGGTGGTGAG 2378
QY	1381	TCGTGCGGCTGAGGTAGCGGTGCTTTCCCGCGCGCTTGGCCCG 1425
Db	2379	TCGTGCGGCTGAGGTAGCGGTGCTTTCCCGCGCGCTTGGCCCG 2423

RESULT 4

AC015998/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC015998 Homo sapiens, clone RP11-15J14, complete sequence. AC015998 Homo sapiens, clone RP11-15J14 HTG. Homo sapiens. Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Birren,B., Linton,L., Nusbaum,C. and Lander,E. 2 (bases 1 to 156873) Homo sapiens, clone RP11-15J14 Unpublished 2 (bases 1 to 156873) Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B., Brown,A., Castle,A., Collangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Llieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye.W.J., Zimmer,A. and Zody,M. Submitted (18-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 3 (bases 1 to 156873) Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRoque,K., Lamazans,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L.,

Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Souneez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (01-MAY-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 21, 2001 this sequence version replaced gi:11560231.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3477

Center clone name: L5_J_14

FEATURES

source	Location/Qualifiers	
	1. .156873	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="RP11-15J14"	
	/clone_lib="RPC1-11 Human Male BAC"	
	1. .87	
repeat_region	/rpt_family="MLT1F1"	/rpt_family="L2"
repeat_region	108. .158	complement(8208. .8688)
repeat_region	/rpt_family="AT_rich"	/rpt_family="L2"
repeat_region	200. .274	8850. .8876
repeat_region	/rpt_family="L2"	/rpt_family="(CATA)n"
repeat_region	1372. .1490	8988. .9097
repeat_region	/rpt_family="MIR"	/rpt_family="L1MB3"
repeat_region	1493. .1539	9098. .9407
repeat_region	/rpt_family="MIR"	/rpt_family="AluJb"
repeat_region	/rpt_family="(TTA)n"	9408. .9498
repeat_region	1599. .1763	/rpt_family="L1MB3"
repeat_region	/rpt_family="FRAM"	9588. .10083
repeat_region	complement(2665. .2732)	/rpt_family="L1MA4A"
repeat_region	2747. .2873	10084. .10129
repeat_region	/rpt_family="L2"	/rpt_family="(CAAAA)n"
repeat_region	3041. .3081	10130. .10892
repeat_region	/rpt_family="L2"	/rpt_family="L1MA4A"
repeat_region	3446. .3472	/rpt_family="L1MA9"
repeat_region	3476. .3501	10893. .11189
repeat_region	/rpt_family="AT_rich"	/rpt_family="AluSx"
repeat_region	3807. .3911	11190. .11413
repeat_region	/rpt_family="MER94"	/rpt_family="L1MA4A"
repeat_region	complement(4180. .4328)	11592. .11845
repeat_region	/rpt_family="L2"	/rpt_family="L1MB3"
repeat_region	complement(5606. .5809)	11886. .11916
repeat_region	/rpt_family="L1P5"	/rpt_family="AT_rich"
repeat_region	5866. .6149	11920. .12153
repeat_region	/rpt_family="AluSx"	/rpt_family="L1MA9"
repeat_region	6332. .6461	12154. .12175
repeat_region	/rpt_family="FLAM_C"	/rpt_family="(TAAA)n"
repeat_region	6492. .6595	12176. .12328
repeat_region	/rpt_family="A-rich"	/rpt_family="L1MA9"
repeat_region	6811. .7150	12329. .12362
repeat_region	/rpt_family="L1MB3"	/rpt_family="(TAAA)n"
repeat_region	7425. .7453	12363. .12670
repeat_region	/rpt_family="AT_rich"	/rpt_family="L1MA9"
repeat_region	7699. .7917	complement(12671. .13025)
repeat_region	complement(7920. .8214)	13026. .13669
		/rpt_family="L1MA9"
		13711. .14022
		/rpt_family="AluSx"
		14082. .14206
		/rpt_family="L1MB3"
		15453. .15532
		/rpt_family="L1MB3"
		16551. .16618
		/rpt_family="L2"
		16730. .16756
		/rpt_family="AT_rich"
		17583. .17891
		/rpt_family="AluSx"
		17814. .17942
		/note="Less than 30 qual SNGL region"
		17901. .18037
		/rpt_family="GA-rich"
		complement(18164. .18466)
		/rpt_family="AluY"
		complement(19025. .19239)
		/rpt_family="MER20"
		19526. .19564
		/rpt_family="(CAAAA)n"
		19917. .19938
		/rpt_family="AT_rich"
		20711. .20821
		/rpt_family="MIR3"
		21514. .22019
		/rpt_family="L2"
		22107. .22209
		/rpt_family="L2"
		complement(22210. .22330)
		/rpt_family="MLT1A1"
		22334. .22989
		/rpt_family="L1"

/evidence-not_experimental
/rpt_family="L2"
9881..9915
/evidence-not_experimental
/rpt_family="(A)n"
9987..10233
/evidence-not_experimental
/rpt_family="MIR"
/complement(10279..10451)
/evidence-not_experimental
/rpt_family="MER20"
10823..10869
/evidence-not_experimental
/rpt_family="L2"
14352..14387
/evidence-not_experimental
/rpt_family="(TCAA)n"
/complement(14742..15115)
/evidence-not_experimental
/rpt_family="MLTlK"
15956..15976
/evidence-not_experimental
/rpt_family="(TC)n"
16580..16624
/evidence-not_experimental
/rpt_family="L2"
17179..17309
/evidence-not_experimental
/rpt_family="MER94"
18483..18503
/evidence-not_experimental
/rpt_family="GC_rich"
18730..18886
/evidence-not_experimental
/rpt_family="MIR"
19782..20109
/evidence-not_experimental
/rpt_family="MLTlJ"
/complement(20367..20418)
/evidence-not_experimental
/rpt_family="MER47"
20423..20482
/evidence-not_experimental
/rpt_family="MLTlJ"
20489..20523
/evidence-not_experimental
/rpt_family="(T)n"
20937..21009
/evidence-not_experimental
/rpt_family="L2"
21112..21141
/evidence-not_experimental
/rpt_family="(TAA)n"
21152..21178
/evidence-not_experimental
/rpt_family="(TTTTC)n"
/complement(21181..21298)
/evidence-not_experimental
/rpt_family="AluJo/FRAM"
21328..21634
/evidence-not_experimental
/rpt_family="MER2"
23969..24028
/evidence-not_experimental
/rpt_family="MIR"
24846..25168
/evidence-not_experimental
/rpt_family="AluSx"
/complement(25610..25702)
/evidence-not_experimental
/rpt_family="L2"
26885..27112
/evidence-not_experimental

repeat_region
/rpt_family="MIR"
29490..29536
/evidence-not_experimental
/rpt_family="(GA)n"
30171..30193
/evidence-not_experimental
/rpt_family="GC_rich"
30215..30269
/evidence-not_experimental
/rpt_family="GC_rich"
31101..31412
/evidence-not_experimental
/rpt_family="L2"
31927..31953
/evidence-not_experimental
/rpt_family="(CA)n"
31953..31979
/evidence-not_experimental
/rpt_family="(GA)n"
/complement(32802..32978)
/evidence-not_experimental
/rpt_family="MIR"
/complement(33702..33866)
/evidence-not_experimental
/rpt_family="MER5A"
34839..34884
/evidence-not_experimental
/rpt_family="LIMA8"
34890..34916
/evidence-not_experimental
/rpt_family="AT_rich"
/complement(34954..35023)
/evidence-not_experimental
/rpt_family="MIR"
/complement(35705..35923)
/evidence-not_experimental
/rpt_family="MIR"
/complement(36531..36718)
/evidence-not_experimental
/rpt_family="MIR"
37008..37226
/evidence-not_experimental
/rpt_family="MIR"
39548..39568
/evidence-not_experimental
/rpt_family="GC_rich"
41600..41706
/evidence-not_experimental
/rpt_family="(TC)n"
42168..42400
/evidence-not_experimental
/rpt_family="MIR"
44440..44511
/evidence-not_experimental
/rpt_family="CT_rich"
45284..45340
/evidence-not_experimental
/rpt_family="GA_rich"
46113..46158
/evidence-not_experimental
/rpt_family="G-rich"
46237..46291
/evidence-not_experimental
/rpt_family="GC_rich"
46918..47077
/evidence-not_experimental
/rpt_family="CT_rich"
/complement(47343..47520)
repeat_region

Query Match 69.5%; Score 991; DB 9; Length 230879;
Best Local Similarity 100.0%; Pred. No. 4e-147;
Matches 991; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 435 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTGATGTGTCCATGCT 494
Db 30771 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTGATGTGTCCATGCT 30712
QY 495 CTCAGACAAGAGAGCTGGTGGCGCGAGCTGGGCTCTTTGCGCCAGCGCCCTCAGC 554
Db 30711 CTCAGACAAGAGAGCTGGTGGCGCGAGCTGGGCTCTTTGCGCCAGCGCCCTCAGC 30652
QY 555 GCCTTGGGGCCACAGCGCGCGCTCCACCTGACAGCTCTCCCTTGGCTTTCGCCCT 614
Db 30651 GCCTTGGGGCCACAGCGCGCGCTCCACCTGACAGCTCTCCCTTGGCTTTCGCCCT 30592
QY 615 ACTGCTGGACGGCGGACCTTGACCCGAGGGGGCGCGCGCGCGCTGGGAAGTCTT 674
Db 30591 ACTGCTGGACGGCGGACCTTGACCCGAGGGGGCGCGCGCGCGCTGGGAAGTCTT 30532
QY 675 CGACGTGTGGCAGGGCTGGCCACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGGGGC 734
Db 30531 CGACGTGTGGCAGGGCTGGCCACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGGGGC 30472
QY 735 CGCATGGGCGAGCTGGACCGCGGGAGGCGCGCGCGCGGGGACCCAGCAACC 794
Db 30471 CGCATGGGCGAGCTGGACCGCGGGAGGCGCGCGCGCGGGGACCCAGCAACC 30412
QY 795 GCGGCCCGCGACCTGCGGAGTCTGGGCTTCGGCGGGAGGGTCCGCCCTCCCCAGGAGCG 854
Db 30411 GCGGCCCGCGACCTGCGGAGTCTGGGCTTCGGCGGGAGGGTCCGCCCTCCCCAGGAGCG 30352
QY 855 GGCCTGTGTGTTATTCACCATGATCCAGCGCAAGAACCTGTTCGCAGAGATGCGCA 914
Db 30351 GGCCTGTGTGTTATTCACCATGATCCAGCGCAAGAACCTGTTCGCAGAGATGCGCA 30292
QY 915 GCAGCTGGGCTCGCGGAGGCTCGGGCCCGCGCGCGCGGGGCTGTGGCCCG 974
Db 30291 GCAGCTGGGCTCGCGGAGGCTCGGGCCCGCGCGCGGGGCTGTGGCCCG 30232
QY 975 GCGCTGGGCGCCCGGATGCGAGGCTTGGTGCCTCGCCCGCGCGCGCGCGCGCG 1034
Db 30231 GCGCTGGGCGCCCGGATGCGAGGCTTGGTGCCTCGCCCGCGCGCGCGCGCG 30172
QY 1035 CAGCGCTTCGCGAGTGCAGTGGCAAGCGGACGCGCAAGAGTCCAGGCTACGCTGCAG 1094
Db 30171 CAGCGCTTCGCGAGTGCAGTGGCAAGCGGACGCGCAAGAGTCCAGGCTACGCTGCAG 30112
QY 1095 CAAGAAGCCCTGCAGCTGAACCTCAAGAGCTGGGCTGGGAGCTGATTCGCGCC 1154
Db 30111 CAAGAAGCCCTGCAGCTGAACCTCAAGAGCTGGGCTGGGAGCTGATTCGCGCC 30052
QY 1155 CCGTGGTACGAGGCTATACCTGCGAGGGTGTATGCGACTTCCCGCTCGCTCGCACT 1214
Db 30051 CCGTGGTACGAGGCTATACCTGCGAGGGTGTATGCGACTTCCCGCTCGCTCGCACT 29992
QY 1215 GGAGCCCAACACCGGCTATCATCAGACGTGATGAATCCATGGACCCCGGCTCCAC 1274
Db 29991 GGAGCCCAACACCGGCTATCATCAGACGTGATGAATCCATGGACCCCGGCTCCAC 29932
QY 1275 CCGCGCCAGCTGCTGGTGCACCAAAATGACTCCCATCAGCATCTATACATCGACGC 1334
Db 29931 CCGCGCCAGCTGCTGGTGCACCAAAATGACTCCCATCAGCATCTATACATCGACGC 29872
QY 1335 GGGCAATATGTTCTACAAGCAGTACGAGACATGTTGGTGGAGTCTGGGCTGCAG 1394
Db 29871 GGGCAATATGTTCTACAAGCAGTACGAGACATGTTGGTGGAGTCTGGGCTGCAG 29812
QY 1395 GTAGCGGTGCTTTCCCGCCGCTTTGGCCG 1425
Db 29811 GTAGCGGTGCTTTCCCGCCGCTTTGGCCG 29781

RESULT 6
162379
LOCUS
DEFINITION Sequence 25 from patent US 5658882.
1171 bp DNA linear PAT 07-OCT-1997

ACCESSION I62379
VERSION I62379.1 GI:2480327
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1171)
AUTHORS Celeste, A. J., Wozney, J. M., Rosen, V. A., Wolfman, N. M., Thomsen, G. H. and Melton, D. A.
TITLE Methods of inducing formation of tendon and/or ligament tissue comprising administering BMP-12, BMP-13, and/or MP-52
JOURNAL Patent: US 5658882-A 25 19-AUG-1997;
FEATURES Location/Qualifiers
Source 1. 1171
BASE COUNT 195 a 384 c 413 g 179 t
ORIGIN
Query Match 69.3%; Score 987; DB 6; Length 1171;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 439 GATCTCTGCACACTCTCTCTCCGGAGACAGAGTATTTGTTGATGTGTCCATGCTCA 498
Db 8 GATCTCTGCACACTCTCTCTCCGGAGACAGAGTATTTGTTGATGTGTCCATGCTCA 67
QY 499 GACAAGAGAGCTGGTGGCGGGAGCTGCGGCTCTTTGCGCAGCGCCCTCAGCGCC 558
Db 68 GACAAGAGAGCTGGTGGCGGGAGCTGCGGCTCTTTGCGCAGCGCCCTCAGCGCC 127
QY 559 TGGGGGCCACACAGCGGGCGCTCCACGTGCAGCTCTTCCCTTGCCTTTCGCCCTACTG 618
Db 128 TGGGGGCCACACAGCGGGCGCTCCACGTGCAGCTCTTCCCTTGCCTTTCGCCCTACTG 187
QY 619 CTGGAGCGCGGACCTTGACCCCGAGGGGGCGCGCGCGCGGCTGGGAAGTCTTCGAC 678
Db 188 CTGGAGCGCGGACCTTGACCCCGAGGGGGCGCGCGCGGCTGGGAAGTCTTCGAC 247
QY 679 GTGTGGCAGGGCTTCGCCACACAGCCCTGGAAGCAGCTGTGTTGGAGCTGGGGCGCA 738
Db 248 GTGTGGCAGGGCTTCGCCACACAGCCCTGGAAGCAGCTGTGTTGGAGCTGGGGCGCA 307
QY 739 TGGGGCGAGCTGGACCGCGGGAGCGCGCGCGCGGGGACCCCGAGCAACCGCGG 798
Db 308 TGGGGCGAGCTGGACCGCGGGAGCGCGCGCGCGGGGACCCCGAGCAACCGCGG 367
QY 799 CCGCGGAGCTCGGAGCTTGGGCTTCGCCCGAGGGTTCGCCCTCCCGAGAGCGGGCC 858
Db 368 CCGCGGAGCTCGGAGCTTGGGCTTCGCCCGAGGGTTCGCCCTCCCGAGAGCGGGCC 427
QY 859 CTGCTGGTGGTATTCACAGATCCAGCGCAAGAACCTGTTCCAGAGATGCGCAGCAG 918
Db 428 CTGCTGGTGGTATTCACAGATCCAGCGCAAGAACCTGTTCCAGAGATGCGCAGCAG 487
QY 919 CTGGCTCGGGCGAGGCTCGGGCGCGCGCGCGGGGCTCGTGGCGCGCGCGG 978
Db 488 CTGGCTCGGGCGAGCTCGGGCGCGCGCGCGGGGCTCGTGGCGCGCGCGG 547
QY 979 TCGGGGGCGCGGATGCGAGGCTTGGCTTCGCCCGCGCGCGCGGGGCGGGCGCAG 1038
Db 548 TCGGGGGCGCGGATGCGAGGCTTGGCTTCGCCCGCGCGCGGGGCGGGCGCAG 607
QY 1039 GCCTTCGCCAGTCGCGATGGCAAGCGGACGCAAGAACCTCAGGCTCAGCTGCAGCAAG 1098
Db 608 GCCTTCGCCAGTCGCGATGGCAAGCGGACGCAAGAACCTCAGGCTCAGCTGCAGCAAG 567
QY 1099 AAGCCCTTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCTTGGGATTCGCGCCCTG 1158
Db 668 AAGCCCTTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCTTGGGATTCGCGCCCTG 727
QY 1159 GAGTACGAGGCTTACCTGCGAGGGTGTATGCGACTTCCCGCTCGGCTCGCACCTGGAG 1218
Db 728 GAGTACGAGGCTTACCTGCGAGGGTGTATGCGACTTCCCGCTCGGCTCGCACCTGGAG 787

QY 1219 CCACCAACACCCATCATCCAGACGCTGATGAATCCATGACCCCGGCTCCACCCCG 1278
 |||||
 Db 788 CCCACCAACACCCATCATCCAGACGCTGATGAATCCATGACCCCGGCTCCACCCCG 847
 |||||
 QY 1279 CCAGCTGCTGCTGCCACCAATATGACTCCCATCAGATTCATATACATCGACGGGGC 1338
 |||||
 Db 848 CCAGCTGCTGCTGCCACCAATATGACTCCCATCAGATTCATATACATCGACGGGGC 907
 |||||
 QY 1339 AATAATGTGTCTACAGAGTACGAGGACATGCTGGTGAGTCGTGGCTCAGGTAG 1398
 |||||
 Db 908 AATAATGTGTCTACAGAGTACGAGGACATGCTGGTGAGTCGTGGCTCAGGTAG 967
 |||||
 QY 1399 CGTGCTTTCCCGCGCCCTTGCCCGC 1425
 |||||
 Db 968 CGTGCTTTCCCGCGCCCTTGCCCGC 994
 |||||

RESULT 7

AC126226

LOCUS

DEFINITION

Papio cynocephalus anubis clone RP41-394G6, linear HTG 04-JUL-2002

3 unordered pieces.

ACCESSION

AC126226

VERSION

HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

olive baboon.

SOURCE

ORGANISM

Papio cynocephalus anubis

Eukaryota; Metazoa; Chordata

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopithecinae; Papio.

1 (bases 1 to 184987)

Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,

Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Harlow, J.,

Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,

Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B.,

Marquillet, E.H., Masello, C., Maskeri, B., Mastrian, S.D.,

McCluskey, J.C., McDowell, J., Paquirigan, C., Pearson, R.,

Portnoy, M.E., Prasad, A., Schueler, M.G., Stantripop, S., Thomas, J.W.,

Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,

Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 184987)

Green, E.D.

Direct Submission

Submitted (04-JUL-2002) NIH Intramural Sequencing Center, 8717

Groveomont Circle, Gaithersburg, MD 20877, USA

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: der

Center clone name: 394G06

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Assembly: Dye-terminator Big Dye; 100% of reads

Consensus quality: 183498 bases at least Q40

Consensus quality: 183965 bases at least Q30

Consensus quality: 184299 bases at least Q20

Insert size: 177000; agarose-fp

Insert size: 18478; sum-of-ctnigs

Quality coverage: 9.40x in Q20 bases; agarose-fp

Quality coverage: 9.00x in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 40078: contig of 40078 bp in length
 * 40079 40178: gap of unknown length
 * 40179 111588: contig of 71410 bp in length
 * 111589 111688: gap of unknown length
 * 111689 184987: contig of 73299 bp in length.

FEATURES

source

1..184987

/organism="Papio cynocephalus anubis"

/db_xref="taxon:9555"

/clone="RP41-394G6"

/clone_lib="RP41"

1..40078

/note="assembly_fragment"

clone_end:SP6

vector_side:right"

40179..111588

/note="assembly_fragment"

111689..184987

/note="assembly_fragment"

clone_end:r7

vector_side:right"

BASE COUNT 52085 a 38650 c 39949 g 54102 t 201 others

ORIGIN

Query Match 65.6%; Score 935; DB 2; Length 184987;
 Best Local Similarity 96.5%; Pred. No. 2.8e-138;
 Matches 956; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 435 AGACGATCTCTCCACACTCTCTCCGAGACAGAGTATTTCTTGTATGTCATGCT 494
 |||||
 Db 178510 AGACGATCTCTCCACACTCTCTCCGAGACAGAGTATTTCTTGTATGTCATGCT 178569

QY 495 CTCAGACAAGAAGAGTGTGGCGCGGAGCTCTTCGCGAGCGCCCTCAGC 554
 |||||
 Db 178570 CTCAGACAAGAAGAGTGTGGCGCGGAGCTCTTCGCGAGCGCCCTCAGC 178629

QY 555 GCCTTGGGGCCACAGCGCGGCGCTCCACGTGACAGCTTCCCTTGTGCTTGC 614
 |||||
 Db 178630 GCCTTGGGGCCACTAGCGGGCGCTCCACGTGACAGCTTCCCTTGTGCTTGC 178689

QY 615 ACTGTGGAGCGCGGACCTTGACCGGAGGCGCGCGCGCGCGGCGGAGTCTT 674
 |||||
 Db 178690 CTTGTGGAGCGCGGACCTTGACCGGAGGCGCGCGCGCGCGGAGTCTT 178749

QY 675 CGACGTGTGGAGGCGCTGCGCCACAGCGCTGGAAGCAGCTGTGTTGAGCTG 734
 |||||
 Db 178750 CGACGTGTGGAGGCGCTGCGCCACAGCGCTGGAAGCAGCTGTGTTGAGCTG 178809

QY 735 CGCATGGGGGAGCTGAGCGCGGGGAGGCGCGCGCGCGCGCGCGCGCGCG 794
 |||||
 Db 178810 CGCATGGGGGAGCTGAGCGCGGGGAGGCGCGCGCGCGCGCGCGCGCG 178869

QY 795 GCCCGCCCGGAGCTGCGGAGTCTGGCTTTCGCGGAGGCTGCGGCTCCCGAG 854
 |||||
 Db 178870 GCCCGCCCGGAGCTGCGGAGTCTGGCTTTCGCGGAGGCTGCGGCTCCCGAG 178929

QY 855 GGCCCTGCTGTGTGTATTCACAGATCCACAGCGCAAGAACTTTCGAGAGATG 914
 |||||
 Db 178930 GGCCCTGCTGTGTGTATTCACAGATCCACAGCGCAAGAACTTTCGAGAGATG 178989

QY 915 GCAGCTGGGCTCGGCGAGGCTCGGGCGCGGGCGCGGGCGCGGGCTGTGTCG 974
 |||||
 Db 178990 GCAGCTGGGCTCGGCGAGGCTCGGGCGCGGGCGCGGGCTGTGTCGTCG 179049

QY 975 GCCGTGGGGCGCGGAGTCCAGAGCTTGGCTGCCCGCGCGCGCGCGCGCG 1034
 |||||
 Db 179050 GCCGTGGGGCGCGGAGTCCAGAGCTTGGCTGCCCGCGCGCGCGCGCGCG 179109

QY 1035 CACGGCTTTCGCCAGTCCGATGCCAAGCGGCGCAGCAAGTCCAGGCTGACG 1094

QY 1283 GCTGCTGGTGGCCACCAAAATTGACTCCATCAGCATCTATACATCGACGGGCAATA 1342
|||||
Db 1196 GCTGCTGGTGGCCACCAAAATTGACTCCATCAGCATCTGTATACATCGACGGGCAATA 1255
|||||
QY 1343 ATGTGGTCTACAAGCAGTACGAGGAGATGGTGGTGGAGTGTGCGGCTGCAGG 1395
|||||
Db 1256 ATGTGGTCTACAAGCAGTACGAGGAGATGGTGGTGGAGTGTGCGGCTGCAGG 1308
|||||

RESULT 10
AC126238 143312 bp DNA linear HTG 04-JUL-2002
LOCUS
DEFINITION
Felis catus clone RP86-410H13, WORKING DRAFT SEQUENCE, 7 unordered
pieces
ACCESSION
AC126238
VERSION
AC126238.1 GI:21693902
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
cat.
ORGANISM
Felis catus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
1 (bases 1 to 143312)
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., McDowell,C., Maskeri,B., Mastrian,S.D.,
McCluskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
NISC
2 (bases 1 to 143312)
Green,E.D.
Direct Submission
Submitted (04-JUL-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: czs
Center clone name: 410H13
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 140814 bases at least Q40
Consensus quality: 141228 bases at least Q30
Consensus quality: 141484 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 142712; sum-of-ctngs
Quality coverage: 9.96x in Q20 bases; agarose-fp
Quality coverage: 9.91x in Q20 bases; sum-of-ctngs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2817: contig of 2817 bp in length
* 2818 2917: gap of unknown length
* 2918 17821: contig of 14904 bp in length
* 17821 17921: gap of unknown length
* 17922 37586: contig of 19665 bp in length

* 37587 37686: gap of unknown length
* 37687 56500: contig of 18814 bp in length
* 56501 56600: gap of unknown length
* 56601 75714: contig of 19114 bp in length
* 75715 75814: gap of unknown length
* 75815 109353: contig of 33539 bp in length
* 109354 109454: gap of unknown length
* 109454 143312: contig of 33859 bp in length.
FEATURES
Location/Qualifiers
1..143312
/organism="Felis catus"
/db_xref="taxon:9685"
/clone="RP86-410H13"
/clone_lib="RP86"
1..2817
/note="assembly_fragment"
2918..17821
/note="assembly_fragment"
17922..37586
/note="assembly_fragment"
37687..56500
/note="assembly_fragment"
clone_end:SP6
vector_side:right
56601..75714
/note="assembly_fragment"
75815..109353
/note="assembly_fragment"
109454..143312
/note="assembly_fragment"
BASE COUNT 39933 a 31312 c 30631 g 40835 t 601 others
ORIGIN
Query Match 62.2%; Score 885.8; DB 2; Length 143312;
Best Local Similarity 94.0%; Pred. No. 1.6e-130;
Matches 933; Conservative 0; Mismatches 57; Indels 3; Gaps 1;
QY 435 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAAATATTTGTGTGATGTGTCATGCT 494
|||||
Db 50385 AGACGATCTCTCGCACACTCTCTCCGGAGACAGAAATATTTGTGTGATGTGTCACACT 50444
|||||
QY 495 CTCAGACAAAGAAGAGCTGGTGGGCGCGAGCTCGGGCTTTTCGCCAGCGCCCTCAGC 554
|||||
Db 50445 CTCAGACAAAGAAGAGCTGGTGGGCGCGAGCTCGGGCTTTTCGCCAGCGCCCAAC 50504
|||||
QY 555 GCCCTGGGGGCCACACAGCGGGCCGCTCCACGTGCAGCTCTCCCTTGCCTTTCGCCCT 614
|||||
Db 50505 GCCCTGGGGGCCGCGCGCGCGCTCCACGTGCAGCTCTTCCCTTGCCTTTCGCCCA 50564
|||||
QY 615 ACTGCTGGACGGCGGACCCCTGGACCCGCGAGGGCGCGCGCGGCTGGGAAGTCTT 674
|||||
Db 50565 GCTGCTGGACGGCGGACCCCTGGACCCGCGAGGGCGCGCGCGGCTGGGAAGTCTT 50624
|||||
QY 675 CGACGTGTGGCAGGGCCCTCGCCACACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGGGGC 734
|||||
Db 50625 CGACGTGTGGCAGGGCCCTCGCCACACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGGGGC 50684
|||||
QY 735 CGCATGGGGCAGCTGGAGCGCGGGAGGCCCGCGCGCGCGCGGGGACC---CCAGCA 791
|||||
Db 50685 CGCGTGGGGCAGCGCGGGAGCTGGGGAGGCCGAGCACGCGCGCGCGCGCGCAGCA 50744
|||||
QY 792 ACCGCGCGCGCGGACCTTGGGAGTCTGGGCTTTCGCGGAGGGTTCGCGCTTCCCGAGCA 851
|||||
Db 50745 GCCACCGCGCGGACCTTGGGAGTCTGGGCTTTCGCGGAGGGTTCGCGCGCGCGCGCAGCA 50804
|||||
QY 852 GCGGGCCCTGCTGTGGTATTACACAGATCCACAGCGCAAGAACCTTGTTCGAGAGATGGC 911
|||||
Db 50805 GCGGCGCCCTGCTGTGCTGTTCACCCAGATCCACAGCGCAAGAACCTTGTTCGCGAGATGGC 50864
|||||
QY 912 CGACGAGCTGGGCTCGGCGCGGAGGCTGCGGGCGCGCGCGCGGGGCTGCTGGCC 971
|||||
Db 50865 CGACGAGCTGGGCTCGGCGCGGAGGCTGCGGGCGCGCGCGCGGGGCTGCTGGCC 50924
|||||

```
QY 972 GCCGCGCTCGGGCGCCCGGATGCCAGGCGCTTGGCTGCCCTCCCGCGCGCGCGCGGCG 1031
Db 50925 GCCGCGCTCGGGCGCCCGGATGCCAGGCGCTTGGCTGCCCTCCCGCGCGCGCGGCG 50984
QY 1032 GCCAGCGGCTTCGCCAGTGCATGCCAGCGGCGGACGCGCAAGAGTCAGGCTAGCGTG 1091
Db 50985 GCCAGCGGCTTCGCCAGTGCATGCCAGCGGCGGACGCGCAAGAGTCAGGCTAGCGTG 51044
QY 1092 CAGCAAGAAGCCCTTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCACTGGATTATCGC 1151
Db 51045 CAGCAAGAAGCCCTTCAGCTGAACCTTCAAGGAGCTGGGCTGGGAGCACTGGATTATCGC 51104
QY 1152 GCCCTGGAGTAGAGGCTTATCACTGCGAGGCTGTATGCGACTTCCCGCTGCGCTCGCA 1211
Db 51105 GCCCTGGAGTAGAGGCTTATCACTGCGAGGCTGTATGCGACTTCCCGCTGCGCTCGCA 51164
QY 1212 CTGTGAGCGCCACCAACAGCCATCATCAGAGCGCTGATGAACCTCCATGAGACCCCGGCTC 1271
Db 51165 CTGTGAGCGCCACCAACAGCCATCATCAGAGCGCTGATGAACCTCCATGAGACCCCGGCTC 51224
QY 1272 CACCCGCGCCAGCTGTGCTGCCGCCACCAAAATGACTCCCATCAGCATTTATACATCGA 1331
Db 51225 CACCCGCGCCAGCTGTGCTGCCGCCACCAAAATGACTCCCATCAGCATTTATACATCGA 51284
QY 1332 CGCGGCAATATGCTGTCTACAGCAGTACGAGGACATGGTGGTGGAGTCGTGCGGCTG 1391
Db 51285 CGCGGCAATATGCTGTCTACAGCAGTACGAGGACATGGTGGTGGAGTCGTGCGGCTG 51344
QY 1392 CAGGTAGCGGTGCTTTCCCGCGCGCTTGCGCC 1424
Db 51345 CAGGTAGCGGTGCTTTCCCGCGCGCTTGCGCC 51377

RESULT 11
AC126232/c
LOCUS
DEFINITION
Bos taurus clone RP42-416E22, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION
AC126232
VERSION
AC126232.1 GI:21693896
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
cow.
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 219978)
AUTHORS
Benjamin,B., Bialesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,
McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 219978)
Green,E.D.
Direct Submission
Submitted (04-JUL-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehgri.nih.gov
----- Project Information
Center project name: djb
Center clone name: 416E22
----- Summary Statistics
```

```
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 213653 bases at least Q40
Consensus quality: 214882 bases at least Q30
Consensus quality: 215481 bases at least Q20
Insert size: 209000; agarose-fp
Insert size: 218878; sum-of-ctnigs
Quality coverage: 7.79x in Q20 bases; agarose-fp
Quality coverage: 7.44x in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3801: contig of 3801 bp in length
* 3802 3901: gap of unknown length
* 3902 6418: contig of 2517 bp in length
* 6419 6518: gap of unknown length
* 6519 12134: contig of 5616 bp in length
* 12135 12235: gap of unknown length
* 12235 18070: contig of 5836 bp in length
* 18071 18170: gap of unknown length
* 18171 37754: contig of 19584 bp in length
* 37755 37854: gap of unknown length
* 37855 57704: contig of 19850 bp in length
* 57705 57805: gap of unknown length
* 57806 84717: contig of 26913 bp in length
* 84718 84817: gap of unknown length
* 84818 108293: contig of 23476 bp in length
* 108294 108393: gap of unknown length
* 108394 130107: contig of 21714 bp in length
* 130108 130207: gap of unknown length
* 130208 155272: contig of 25065 bp in length
* 155273 155372: gap of unknown length
* 155373 183289: contig of 27917 bp in length
* 183290 183389: gap of unknown length
* 183390 219978: contig of 36589 bp in length.
*
* Location/Qualifiers
* 1. 219978
* /organism="Bos taurus"
* /db_xref="taxon:9913"
* /clone="RP42-416E22"
* /clone_lib="RP42"
* 1. 3801
* /note="assembly_fragment"
* 3902. 6418
* /note="assembly_fragment"
* 6519. 12134
* /note="assembly_fragment"
* 12235. 18070
* /note="assembly_fragment"
* 18171. 37754
* /note="assembly_fragment"
* 37855. 57704
* /note="assembly_fragment"
* clone_end:SP6
* vector_side:right"
* 57805. 84717
* /note="assembly_fragment"
* 84818. 108293
* /note="assembly_fragment"
* 108394. 130107
* /note="assembly_fragment"
* 130208. 155272
* /note="assembly_fragment"
* 155373. 183289
* /note="assembly_fragment"
* 183390. 219978
```

```
/note="assembly_fragment"
BASE COUNT 59915 a 49896 c 49290 g 59765 t 1112 others
ORIGIN

Query Match 56.1%; Score 798.8; DB 2; Length 219978;
Best Local Similarity 89.7%; Pred. No. 7.7e-111;
Matches 913; Conservative 0; Mismatches 77; Indels 28; Gaps 4;

QY 435 AGACGATCTCTCGCACACCTCTCTCCGGAGACAGAGTAGTATTTGTTTGATGTGTCACATGCT 494
|||||
Db 126101 AGACGATCTCTCGCACACCTCTCTCCGGAGACAGAGTAGTATTTGTTTGATGTGTCACAGCT 126042

QY 495 CTCAGACAAAGAGAGCTGGTGGGCGCGAGCTCGGGCTCTTTTCGCCAGCGCCCTCAGC 554
|||||
Db 126041 CTCAGACAAAGAGAGCTGGTGGGCGCGAGCTCGGGCTCTTTTCGCCAGCGCCCTCAGC 125982

QY 555 GCCCTGGGGGCCACACGCGGGCGGCTCCACGTGACGTCTTCCCTTGGCTTTCGCCCTT 614
|||||
Db 125981 GCCCTGGGGGCCCGCGCGGGCGGCTCCGCTTGCAAGCTCTTTCGCCCTTTCGCCCTT 125922

QY 615 ACTGCTGGACGCGCGGACCTTGACCCGAGGGGCGCGCGGCTGGGAGTCTT 674
|||||
Db 125921 GCTGCTGGAAGCGGGAGCTTGACCCGAGGGGCGCGCGGCTGGGAGTCTT 125862

QY 675 CGACGTGTGGCAGGGCTCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGCTGCGGGC 734
|||||
Db 125861 CGACGTGTGGCAGGGCTCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGCTTGGGC 125802

QY 735 CGCATGG---GCGAGCTGGAGCGGGGAGCGCGGCGCGGCGGGGACCCAGCA 791
|||||
Db 125801 CGCGTGGGCGGCGGCGGGGCGCGGAGGACGAGCGCGCGGCTGGGCGCCAGCA 125742

QY 792 ACCGCGCGCCCGGACCTCGGAGTGTGGCTTCGGCGGAGGGTGGCGCTCCCGAGCA 851
|||||
Db 125741 GCGCGCGCCCGGACCTCGGAGTGTGGCTTCGGCAGAGGGTGGCGACCCCGAGCA 125682

QY 852 GCGGGCCCTGCTGGTGTATTACACAGATCCAGCGCAAGAACCTGTTCCGACAGATGCG 911
|||||
Db 125681 GCGCGCTTGTCTGCTGTGTTCCAGTCCAGGCAAGACCTGTTCCGCGAGATGCG 125622

QY 912 CGAGCAGCTGGGCTGG---CGAGGCTCGGGCGCGGCGCGGCGGCGGAGGCTGCTG 968
|||||
Db 125621 CGAGCAGCTGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGAGGCTGCGG 125562

QY 969 -----GCCGCGCGCTCGGGCGCGCGCGGCTTCAGGCGCTTGGCT 1007
|||||
Db 125561 GCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGGCGCGCGGCTTGGCT 125502

QY 1008 GCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCGAGTTCGCCATGCGAAGCGCA 1067
|||||
Db 125501 GCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCGAGTTCGCCATGCGAAGCGCA 125442

QY 1068 CGGCAAGAGTCCAGGCTACGCTGACGAGAGAGCCCTGCGAGCTGAAGTTCAGAGGCT 1127
|||||
Db 125441 CGGCAAGAGTCCAGGCTACGCTGACGAGAGAGCCCTGCGAGTGAAGTTCAGAGGCT 125382

QY 1128 GGGCTGGGACGACTGGATTATCGCGCCCTGGAGTACGAGGCTTATCATCTGCGAGGCTG 1187
|||||
Db 125381 GGGCTGGGACGACTGGATTATCGCGCCCTGGAGTACGAGGCTTATCATCTGCGAGGCTG 125322

QY 1188 ATGCGACTTCCCGCTGCGCTGCGACCTGAGCGCCACCAACACGCGCATCATCCAGCGT 1247
|||||
Db 125321 GTGCGACTTCCCGCTGCGCTGCGACCTGAGCGCCACCAACACGCGCATCATCCAGCGT 125262

QY 1248 GATGAATCCATGACCGCGGCTCCACCGCGCGGCTGCTGCGCGCCCAAAATTGAC 1307
|||||
Db 125261 GATGAATCCATGACCGCGGCTCCACCGCGCGGCTGCTGCGCGCCCAAAATTGAC 125202

QY 1308 TCCCATCAGCATCTTATACATCAGCGCGGCAATAATGTGGTCTACAAGCAGTACGAGGA 1367
|||||
Db 125201 TCCCATCAGCATCTTATACATCAGCGCGGCAATAATGTGGTCTACAAGCAGTACGAGGA 125142

QY 1368 CATGTGTGGAGTGTGCGGGCTGCGAGGTAGCGGTGCGCTTTCCCGCGCGCTTGGCGCG 1425
|||||
Db 125141 GATGTGTGGAGTGTGCGGGCTGCGAGGTAGCGGTGCGCTTTCCCGCGCGCTTGGCGCG 125085

|||||
Db 125141 GATGTGTGGAGTGTGCGGGCTGCGAGGTAGCGGTGCGCTTTCCCGCGCGCTTGGCGCG 125085

RESULT 12
AC058786/c
LOCUS 209733 bp DNA linear ROD 29-JAN-2002
DEFINITION Mus musculus clone rp23-117o7, complete sequence.
ACCESSION AC058786
VERSION AC058786.23 GI:11120825
KEYWORDS HTG.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 209733)
AUTHORS Pan, H. and Roe, B.A.
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL Mus musculus BAC Clone rp23-117o7
REFERENCE 2 (bases 1 to 209733)
AUTHORS Pan, H. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE 3 (bases 1 to 209733)
AUTHORS Pan, H. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2000) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
REFERENCE 4 (bases 1 to 209733)
AUTHORS Pan, H. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT On Nov 8, 2000 this sequence version replaced gi:11119463.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source
location/Qualifiers
1..209733
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="rp23-117o7"
/culture="RP23"
BASE COUNT 61917 a 45381 c 44340 g 58095 t
ORIGIN

Query Match 53.7%; Score 765.8; DB 10; Length 209733;
Best Local Similarity 87.4%; Pred. No. 1.2e-111;
Matches 866; Conservative 0; Mismatches 112; Indels 13; Gaps 2;

QY 435 ACACGATCTCTCGCACACCTCTCTCCGGAGACAGAGTAGTATTTGTTTGATGTGTCACATGCT 494
|||||
Db 55214 AGACGATCTCTCGCACACCTCTCTCCGGAGACAGAGTAGTATTTGTTTGATGTGTCACACT 65155

QY 495 CTCAGACAAAGAGAGCTGGTGGGCGCGGAGCTGCGGCTCTTTTCGCCAGCGCCCTCAGC 554
|||||
Db 65154 CTCAGACAAAGAGAGCTGGTGGGCGCGGAGCTAAGGCTTTATCGCAGCGCGCCCAAC 65095

QY 555 GCCCTGGGGGCCACACGCGGGCGGCTCCACGTGACGTCTTCCCTTGGCTTTCGCCCT 614
|||||
Db 65094 GCCCTGGGGGCCACACGCGGGCGGCTCCACGTGACGTCTTCCCTTGGCTTTCGCCCT 65035

QY 615 ACTGCTGGACGCGCGGACCTTGACCCGAGGGGCGCGCGGCTGGGAGTCTT 674
|||||
Db 65034 GCTACTGGACGCGGACCTTGATCCTCAGGAGCAACCCAGCGCGCTTGGAGTCTT 64975

QY 675 CGACGTGTGGCAGGGCTCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGCTGCGGGC 734
|||||
Db 675 CGACGTGTGGCAGGGCTCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGCTGCGGGC 734
```

Db	64974	CGACGTGTGCGAGGGCTCGCCCTCAGCCTTGGAAAGACGTGTGCTGTGAGTTGGCGG	64915
QY	735	CGACTGGGGAGCTGACGCCGGGAGGCGGAGCGCGCGCGGACCCAGCAACC	794
Db	64914	AGCTTGGGTGAGCTGAGCGCCGGGATAGCGGGGCGCGCGAGGGGTCTCCACGAGCG	64855
QY	795	GCCGCCCCGGAGCTCGGAGTCTGGGCTTCGCCGCGAGGCTCGCCCTCCCGAGGAGCG	854
Db	64854	ACCGCTCTGGACCTCGGAGTCTGGCTTCGGTCGAGGGTGAAGCGCCCGCCAGGAGCG	64795
QY	855	GGCCCTGCTGGTGGTATTACACAGATCCAGCGCCAGAACCTTTTCGAGAGATGCGCGA	914
Db	64794	CGCCCTGCTTGTAGTCTTACACAGATCGAGCGCAAGAACCTTTCTACTGAGATGATGA	64735
QY	915	GCAGCTGGGCTCGCGAGGCTCGCGGCCCGCGCGCGCGCGGCGGCTGTCGGCGCC	974
Db	64734	GCAGCTGGGCTCTGCAGAGCT-----GCGGAGCGAGGGGTTCATGGCCAGC	64687
QY	975	GCCGTGCGGGCGCCCGGATGCCAGGCTTGGCTGCCCTCGCCCGCGCGCGCGCGGCGG	1034
Db	64686	GCCGTGCGGGCTCCCGAGCGCGGGTCTTGGCTGCCCTCGCCCGCGCGCGCGCGGCG	64627
QY	1035	CAGCGCTTCGCCAGTCGCATGCGCAAGCGGCACGCGCAAGATCCAGGCTACGCTGCAG	1094
Db	64626	CACCGCTTCGCCAGCGCTCACGCGCAAGCGACATGGCAAGTCCAGGCTGCGCTGCAG	64567
QY	1095	CAAGAAGCCCTCGACGTGAACCTTCAAGGAGCTGGGCTGGGAGGACTGATATTCGCGC	1154
Db	64566	CAGAAGCCCTCGACGTGAATTTAAGAGTTAGCTGGGAGGACTGATATTCGCGC	64507
QY	1155	CTGGAGTACGAGGCTATCACTGCGAGGGTGTATCGACTTCCCGCTGCGCTCGCACCT	1214
Db	64506	CCTAGAGTACGAGGCTATCACTGCGAGGGGTGTCCGACTTTCGCTGCGCTCGCACCT	64447
QY	1215	GGAGCCCAACCAACCGCATCATCAGAGCTGATGAAGTCCATGAGCCCGGCTCCAC	1274
Db	64446	TGAGCCCAATGAACCTATCACTATTCAGACGCTGATGAATCCATGAGCCCGGCTCCAC	64387
QY	1275	CCGCGCCAGCTGCTGGTGGCCACCAATTTGACTCCCATCAGCATCTATACATCGACGC	1334
Db	64386	CCGCGCTAGCTGCTGGTGGCCACCAATTTGACTCCCATCAGCATCTATACATCGACGC	64327
QY	1335	GGCAATATGTTGTACAGCAGTACGAGGACATGGTGGTGGAGTCTGCGGCTGCAG	1394
Db	64326	GGCAATATGTTGTACAGCAGTATGAGGACATGGTGGTGGAGTCTGCGGCTGTAG	64267
QY	1395	GTAGCGGTGCTTCCCGCGGCTTGGCCG 1425	
Db	64266	GTAGCGGTG-CTGTCCCGCACCTGGGCCAG 64237	
RESULT	13		
AL732476			
LOCUS	AL732476	326540 bp DNA linear HTG 25-JUN-2002	
DEFINITION	Mus musculus chromosome 4 clone RP23-131A6, *** SEQUENCING IN		
PROGRESS	***, 13 unordered pieces.		
ACCESSION	AL732476		
VERSION	AL732476		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 326540)		
JOURNAL	Sycamore,N.		
COMMENT	Direct Submission Submitted (24-JUN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jun 26, 2002 this sequence version replaced gi:21531590. ----- Genome Center Center: Wellcome Trust Sanger Institute		

Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM131A6
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 320895 bases at least Q40
Consensus quality: 323247 bases at least Q30
Consensus quality: 324562 bases at least Q20
Insert size: 325340; sum-of-contigs
Insert size: 175863; 3.3% error; agarose-fp
Quality coverage: 5.99x in Q20 bases; sum-of-contigs Quality
coverage: 11.18x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 15978: contig of 15978 bp in length
* 15979 16078: gap of 100 bp
* 16079 48717: contig of 32639 bp in length
* 48718 48817: gap of 100 bp
* 48818 92124: contig of 43307 bp in length
* 92125 92224: gap of 100 bp
* 92225 127311: contig of 35087 bp in length
* 127312 127411: gap of 100 bp
* 127412 159291: contig of 31880 bp in length
* 159292 159391: gap of 100 bp
* 159392 182357: contig of 22966 bp in length
* 182358 182458: gap of 101 bp
* 182459 273966: contig of 91508 bp in length
* 273967 274067: gap of 101 bp
* 274068 283037: contig of 8970 bp in length
* 283038 283137: gap of 100 bp
* 283138 292483: contig of 9346 bp in length
* 292484 292583: gap of 100 bp
* 292584 295787: contig of 3204 bp in length
* 295788 295887: gap of 100 bp
* 295888 302884: contig of 6997 bp in length
* 302885 302984: gap of 100 bp
* 302985 321370: contig of 18386 bp in length
* 321371 321470: gap of 100 bp
* 321471 326540: contig of 5070 bp in length.

FEATURES
Location/Qualifiers
1..326540
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-131A6"
/clone_lib="RPCI-23"
1..15978
/note="assembly_fragment:03237
fragment_chain:1"
16079..48717
/note="assembly_fragment:03208
fragment_chain:1"
48818..92124
/note="assembly_fragment:02621
fragment_chain:1"
92225..127311
/note="assembly_fragment:03198
fragment_chain:1"
127412..159291
/note="assembly_fragment:02086
fragment_chain:1"
159392..182357
/note="assembly_fragment:01805
fragment_chain:1"


```
fragment_chain:1
clone_end:77
vector_side:left
182458..273966
/note="assembly_fragment:00350"
274067..283037
/note="assembly_fragment:00833.0"
283138..292483
/note="assembly_fragment:02119"
292584..295787
/note="assembly_fragment:02468"
295888..302884
/note="assembly_fragment:02675"
302985..321370
/note="assembly_fragment:02846"
321471..326540
/note="assembly_fragment:02945"
BASE COUNT 93482 a 65929 c 68488 g 97422 t 1219 others
ORIGIN
Query Match 53.0%; Score 755.4; DB 2; Length 326540;
Best Local Similarity 87.4%; Pred No.5e-110;
Matches 866; Conservative 0; Mismatches 111; Indels 14; Gaps 3;
QY 435 AGACGATCTCTCGACACTCTCTCGGAGACAGAAAGTATTTTGTGATGTGTCCACTGT 494
Db 31033 AGACGATCTCTCGACACTCTCTCGGAGACAGAAAGTATTTTGTGATGTGTCCACTGT 31092
QY 495 CTCAGACAAGAGAGCTGGTGGGCGGAGCTGGGCTCTTCGCCAGGCGCCCTCAGC 554
Db 31093 CTCAGACAAGAGAGCTGGTGGGCGGAGCTGAAGCTTTATCGCCAGGCGCCCCCAAC 31152
QY 555 GCCTGGGGCCACAGCGGGCGCTGCACGCTGCAGCTTCCTCCCTTCTTGGCCCT 614
Db 31153 GCCTGGGGCTACCGGCCGACCGCTGCAGCTTCCTCCCTTGTATCCCAT 31212
QY 615 ACTGTGGACGCGCGACCTCTGACCCGAGGGGGCGCGCGCGCGCTGGGAAGTCTT 674
Db 31213 GCTACTGGACGCGGACCTCTGATCTCAGGAGCAACCCAGCGCGCTGGGAAGTCTT 31272
QY 675 CGAGCTGTGGCAGGCGCTCGCCACAGCCCTTGGAGAGAGCTGTGCTTGGAGCTGGGGC 734
Db 31273 CGAGCTGTGGCAGGCGCTCGCCCTCAGCTTGGAGAGAGCTGTGCTTGGAGCTGGGGC 31332
QY 735 CGCATTTGGCGAGCTGGAGCCGCGGAGCGCGCGCGCGCGCGCCAGCAAC 794
Db 31333 AGCTGGGTGAGCTGGAGCGCGGGGATACGGGGCGCGCGGAGGGTCCCGAGCAGCC 31392
QY 795 GCGGCCCCGGACCTCGGAGCTGTGGGCTTCGGCGGAGGGTGGGCGCTTCCCGAGGAGCG 854
Db 31393 ACCGCTCTGGACCTCGGAGCTGTGGGCTTCGGGTTTCGGAGGGTGAGGCGCCCGAGGAGCG 31452
QY 855 GGCCTCTGTGTGATTTACAGATCCAGCGCAAGAACCTTTCGCGAGAGATGGCGCA 914
Db 31453 CGCCTCTGTGTGATTTACAGATTCGAGCGCAAGAACCTTTCCTACTGAGATGCATGA 31512
QY 915 GCAGCTGGCTCGGCGAGGCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 974
Db 31513 GCAGCTGGCTCTGACAGGCT-----CGGGAGCGCGGGGTCTATGGCCAGC 31560
QY 975 GCGCTGGGCGCGCGGATGCGAGCGCTTGGCTGCCCTCCCGCGCGCGCGCGCGCGCGCG 1034
Db 31561 GCGCTGGGCTCCCGAGCGCGGGTCTTGGCTGCCCTCCCGCGCGCGCGCGCGCGCGCG 31620
QY 1035 CAGGCGCTTCGCGAGTGGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1094
Db 31621 CAGGCGCTTCGCGAGCGCTGCGGCAAGCGCA-GGCAAGAGTCCAGGCTCGCTGCGAG 31679
QY 1095 CAAGAAGCCCTGCGAGCTGAAGAGCTGGGCTGGGAGAGCTGGGATTCGCGCC 1154
Db 31680 CAGAAGCCCTGCGAGCTGAAGAGCTGGGATTCGCGCC 31739
QY 1155 CTTGAGTACGAGGCGCTATCTAGTGGAGGGTGTATGCGACTTCCGCTCGCCTCGACCT 1214
```

```
Db 31740 CCTAGTAGTAGGCGCTATCACTGCGAGGCGGTGTCGACTTTCGCGCTCGCACCT 31799
QY 1215 GGAGCCCCAACACAGCGCATCATCCAGACGCTGATGAAGTCCATGAGCCCGGCTCCAC 1274
Db 31800 TGAGCCCACTAAACATCCATCATTCAGACGCTGATGAAGTCCATGAGCCCGGCTCCAC 31859
QY 1275 CCGGCCAGCTGCTGCTGCCACCAAAATTGACTCCCATCAGCATCTATATACATCAGCGC 1334
Db 31860 CCGCCTAGCTGCTGCTGCCACCAAACTGACTCCCATTAGCATCTGTATATCAGCGC 31919
QY 1335 GGSCAATAATGTGCTTACAAGCAGTACGAGGACATGGTGGTGGAGTGTGCGGCTGCAG 1394
Db 31920 GGSCAATAATGTAGTCTACAAGCAGTATGAGGACATGGTGGTGGAGTGTGCGGCTGTAG 31979
QY 1395 GTAGCGGTGCTTTCGCGCGCGCTTGGCGCG 1425
Db 31980 GTAGCGGTG-CTGTCCCGCACCTGGGCGCAG 32009

RESULT 14
AC116927
LOCUS
DEFINITION
AC116927 167312 bp DNA linear HTG 03-APR-2002
DRAFT SEQUENCE, 4 unordered pieces.
AC116927
AC116927.1 GI:19909401
HTG: HTGS_PHASE1; HTGS_DRAFT.
Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 167312)
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Marquies,E.H., Masiello,C., Maskeri,B., Mastrian,S.D.,
McCluskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantripop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 167312)
Green,E.D.
Direct Submission
Submitted (03-APR-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nihri.nih.gov
----- Project Information
Center project name: cxs
Center clone name: 496H14
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 166108 bases at least Q40
Consensus quality: 166132 bases at least Q30
Consensus quality: 166493 bases at least Q20
Insert size: 192000; agarose-fp
Quality coverage: 167012; sum-of-contigs
Quality coverage: 10.34x in Q20 bases; agarose-fp
Quality coverage: 11.89x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
consists of 4 contigs. The true order of the pieces
```


* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 4963: contig of 4963 bp in length
* 4964 5063: gap of unknown length
* 5064 38133: contig of 33070 bp in length
* 38134 38233: gap of unknown length
* 38234 91212: contig of 52979 bp in length
* 91213 91312: gap of unknown length
* 91313 167312: contig of 76000 bp in length.

FEATURES

source
1. .167312
/organism="Rattus norvegicus"
/strain="Brown Norway"
/db_xref="taxon:10116"
/clone_lib="RP31"
/clone_lib="RP31"
1. .4963
/note="assembly_fragment"
clone_end:SP6
vector_side:right
5064. .38133
/note="assembly_fragment"
clone_end:T7
vector_side:right
38234. .91212
/note="assembly_fragment"
91313. .167312
/note="assembly_fragment"
BASE COUNT 47363 a 35040 c 36627 g 47981 t 301 others
ORIGIN

Query Match 52.6%; Score 749.8; DB 2; Length 167312;
Best Local Similarity 86.4%; Pred. No. 4.2e-109;
Matches 856; Conservative 0; Mismatches 122; Indels 13; Gaps 2;

QY 435 AGAGATCTTCGCACACTCTCTCCGAGACAGAGTATTGTTGATGTGCCATGCT 494
|||||
Db 76193 AGAGATCTTCGCACACTCTCTCCGAGACAGAGTATTGTTGATGTGCCATGCT 76252
|||||

QY 495 CTCAGACAAAGAGCTGTGGCGCGGAGCTGCGCTCTTTGCCAGCGCCCTCAGC 554
|||||
Db 76253 CTCAGACAAAGAGCTGTGGCGCGGAGCTGAGCTTTATCGCCAGCGCCCAAC 76312
|||||

QY 555 GCCTGGGGCCACCAGCGCGGCTCCAGCTGCAGCTTCCCTTGCTTGCCTTGC 614
|||||
Db 76313 GCCTGGGGCCCAACACCGCAGCGCTGCATTTGCAGCTTCCCTTGTTGCCCAT 76372
|||||

QY 615 ACTGCTGGACGCGGACCTTGACCGCAGGGCGCGCGCGCGCTGGGAAGTCTT 674
|||||
Db 76373 GTTACTGGACTTAGGACCTTGATCTCAGGGGCCAACAGAGCTGGCTGGGAAGTTT 76432
|||||

QY 675 CGAGTGTGCGAGGGCTGCGCCACCAGCCCTGGAAGCAGCTGTGCTGGAGCTGCGGGC 734
|||||
Db 76433 CGAGTGTGCGAGTCTGCGCCCTCAGCGCTGGAAGCAGCTGTGCTGGAGCTTGGGC 76492
|||||

QY 735 CGATGGGGCGAGCTGGACCGCGGGAGCGGAGCGCGCGCGGGGACCCAGCAACC 794
|||||
Db 76493 AGTCTGGGGTGAGCTGGACCGCCAGGGATTACAGGAGCGCGCGCGGGGTCCCAACAGTC 76552
|||||

QY 795 GCGCGCCCGGACCTGGGAGTCTGGGCTTCGCGCGGAGGGTGGCGCTCCCGCAGGAGCG 854
|||||
Db 76553 ACCGCGCTTGAGCTGCGGAGTCTGGGCTTCGCGCGGAGGGTGGCGCGCGCGGAGCG 76612
|||||

QY 855 GGCCCTGCTGGTGGTATTCACAGATCCCGCAGCGCAAGAACCTGTTCGACAGATGCGCGA 914
|||||
Db 76613 CGCCCTGCTGTGCTTACAGATCGCAGCGCAAGAACCTGTTCACCTGAGATGATGA 76672
|||||

QY 915 GCAGCTGGGCTCGCGCAGAGCTCGGGGCGCGGGCGCGCGGGGCTGTGTGGCGGC 974
|||||

Db 76673 GCAGCTGGGGCTCTGCAGAGGCT-----GCCGGAGCCGAGGGGTCTATGGCCAGC 76720
QY 975 GCCTGCGGGCGCCCGGATCCAGGCTTGGCTGCCCTCCCGCGCGCGCGCGCGCGCG 1034
|||||
Db 76721 GCCTGCGGGCGCCCGGATCCAGGCTTGGCTGCCCTCCCGCGCGCGCGCGCGCG 76780
|||||

QY 1035 CAGGGCTTCCGAGTGCCTATGCGGAGCGGCAAGTCCAGGCTACGCTCGAC 1094
|||||
Db 76781 CAGCGCTTCCGAGTGCCTATGCGGAGCGGCAAGTCCAGGCTACGCTCGAC 76840
|||||

QY 1095 CAAGAGCCCTTCCGAGTGCCTATGCGGAGCGGCAAGTCCAGGCTACGCTCGAC 1154
|||||
Db 76841 CAGAAACCTTCCGAGTGCCTATGCGGAGCGGCAAGTCCAGGCTACGCTCGAC 76900
|||||

QY 1155 CTGGAGTACGAGGCTTATCACTGCGAGGCTGTATGCACTTCCCGTGGCTCGACCT 1214
|||||
Db 76901 CTTAGAGTACGAGGCTTATCACTGCGAGGCGGTGCGGACTTTCGCTGCGACCT 76960
|||||

QY 1215 GGAGCCCAACACGCTATCACTGCGAGCGGTGATGAACTCCATGACCCCGGCTCCAC 1274
|||||
Db 76961 GGAGCCCAACACGCTATCACTGCGAGCGGTGATGAACTCCATGACCCCGGCTCCAC 77020
|||||

QY 1275 CCGCCAGCTGCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1334
|||||
Db 77021 CCCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 77080
|||||

QY 1335 GGGCAATAATGTGCTTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1394
|||||
Db 77081 GGGCAATAATGTGCTTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 77140
|||||

QY 1395 GTAGCGGTGCTTTCGCGCGCGCTTGGCGCG 1425
|||||

Db 77141 GTAGCGGTG-CTGTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 77170

RESULT 15
162382
LOCUS
DEFINITION Sequence 31 from patent US 5658882.
ACCESSION 162382
VERSION 162382.1 GI:2480330
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1046)
AUTHORS Celeste,A.J., Wozney,J.M., Rosen,V.A., Wolfman,N.M., Thomsen,G.H.
and Melton,D.A.
TITLE Methods of inducing formation of tendon and/or ligament tissue
comprising administering BMP-12, BMP-13, and/or MP-52
JOURNAL Patent: US 5658882-A 31 19-AUG-1997;
FEATURES Location/Qualifiers
source
1. .1046
BASE COUNT 212 a 314 c 331 g 187 t 2 others
ORIGIN

Query Match 40.8%; Score 581.8; DB 6; Length 1046;
Best Local Similarity 87.2%; Pred. No. 2.8e-82;
Matches 675; Conservative 1; Mismatches 84; Indels 14; Gaps 3;

QY 653 CGCGCGCGCGCTGGGAAGTC-TTCGAGCTGTGGCAGGCGCTTGGGCCACACCGCTTGAAG 711
|||||
Db 59 CCCAGCGCGCTGGGAAGTCTTTCGAGCTGTGGCAGGCGCTTGGGCCACACCGCTTGAAG 118
|||||

QY 712 CAGCTGTGCTGGAGCTGCGGGCGCGCATGCGGAGCTGAGCGCGCGGGAGCGAGCGG 771
|||||
Db 119 CAGCTGTGCTGGAGTGTGCGGCGAGCTGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAG 178
|||||

QY 772 CGCGCGGGGACCCAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 831
|||||

Db 179 CGCGCGAGGGGTCCCCAGCAGCACCGCGCTTGGACCTCGCGAGTCTTGGGCTTGGGTCG 238
|||||

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 01:01:20 ; Search time 78 Seconds
(without alignments)
5602.749 Million cell updates/sec

Title: US-09-825-751A-19
Perfect score: 1425
Sequence: 1 ctctctgggagacgcagccca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987	69.3	1171	1	US-08-362-670B-25
2	987	69.3	1171	3	US-08-333-576C-25
3	987	69.3	1171	4	US-08-808-324-25
4	987	69.3	1171	5	PCT-US94-14030A-25
5	581.8	40.8	1046	1	US-08-362-670B-31
6	581.8	40.8	1046	3	US-08-333-576C-31
7	581.8	40.8	1046	4	US-08-808-324-31
8	581.8	40.8	1046	5	PCT-US94-14030A-31
9	412.2	28.9	530	1	US-08-581-529B-5
10	412.2	28.9	530	3	US-09-097-616-5
11	412.2	28.9	530	5	PCT-US94-07762-5
12	259.2	18.2	1345	1	US-08-362-670B-33
13	259.2	18.2	1345	3	US-08-333-576C-33
14	259.2	18.2	1345	4	US-08-808-324-33
15	259.2	18.2	1345	5	PCT-US94-14030A-33
16	253	17.8	2703	2	US-08-288-508C-1
17	253	17.8	2703	4	US-09-180-700-4
18	251.6	17.7	1207	1	US-08-362-670B-3
19	251.6	17.7	1207	3	US-08-333-576C-3
20	251.6	17.7	1207	5	US-08-289-222E-1
21	251.6	17.7	1207	4	US-09-054-526B-1
22	251.6	17.7	1207	4	US-08-808-324-3
23	251.6	17.7	1207	5	PCT-US94-14030A-3
24	249.2	17.5	519	2	US-08-581-528A-5
25	249.2	17.5	519	5	PCT-US94-07799-5
26	245.8	17.2	357	4	US-09-180-700-1
27	245	17.2	1203	1	US-08-362-670B-29

28	245	17.2	1203	3	US-08-333-576C-29	Sequence 29, Appl
29	245	17.2	1203	4	US-08-808-324-29	Sequence 29, Appl
30	245	17.2	1203	5	PCT-US94-14030A-29	Sequence 29, Appl
31	244.8	17.2	926	1	US-08-362-670B-1	Sequence 1, Appl
32	244.8	17.2	926	3	US-08-333-576C-1	Sequence 1, Appl
33	244.8	17.2	926	4	US-08-808-324-1	Sequence 1, Appl
34	244.8	17.2	926	5	PCT-US94-14030A-1	Sequence 1, Appl
35	239	16.8	2329	1	US-08-455-559-9	Sequence 9, Appl
36	239	16.8	2329	4	US-09-145-060-9	Sequence 9, Appl
37	239	16.8	2329	5	PCT-US94-00657-9	Sequence 9, Appl
38	232	16.3	272	1	US-08-362-670B-12	Sequence 12, Appl
39	232	16.3	272	3	US-08-333-576C-12	Sequence 12, Appl
40	232	16.3	272	4	US-08-808-324-12	Sequence 12, Appl
41	232	16.3	272	5	PCT-US94-14030A-12	Sequence 12, Appl
42	230	16.1	1233	1	US-08-362-670B-27	Sequence 27, Appl
43	230	16.1	1233	3	US-08-333-576C-27	Sequence 27, Appl
44	230	16.1	1233	4	US-08-808-324-27	Sequence 27, Appl
45	230	16.1	1233	5	PCT-US94-14030A-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-362-670B-25
; Sequence 25, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Mooney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Human VL-1 protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..964
; FEATURE:
; NAME/KEY: mat_peptide

LOCATION: 605..964
US-08-362-670B-25

Query Match 69.3%; Score 987; DB 1; Length 1171;
Best Local Similarity 100.0%; Pred. No. 2.7e-185;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCTCATGCTCTCA 498
|||||
Db 8 GATCTCTCGCACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCTCATGCTCTCA 67
|||||

QY 499 GACAAAGAAGAGCTGTTGGCGCGGAGCTGCGGCTCTTTTCGCCAGGCGCCCTCAGGCGCC 558
|||||
Db 68 GACAAAGAAGAGCTGTTGGCGCGGAGCTGCGGCTCTTTTCGCCAGGCGCCCTCAGGCGCC 127
|||||

QY 559 TGGGGGCCACACCGCGCGCTCCACGTGACGCTCTTCCTTTGCTTTCGCCCTTACTG 618
|||||

Db 128 TGGGGGCCACACCGCGCGCTCCACGTGACGCTCTTCCTTTGCTTTCGCCCTTACTG 187
|||||

QY 619 CTGGACGCGGACCTTGACCGCGAGGGCGCGCGCGGCTGGGAAGTCTTCGAC 678
|||||

Db 188 CTGGACGCGGACCTTGACCGCGAGGGCGCGCGCGGCTGGGAAGTCTTCGAC 247
|||||

QY 679 GTGTGCGAGGCTGCGCCACCAAGCCTTGAAGACGCTGTGCTTGAGCTGCGGCGCGCA 738
|||||

Db 248 GTGTGCGAGGCTGCGCCACCAAGCCTTGAAGACGCTGTGCTTGAGCTGCGGCGCGCA 307
|||||

QY 739 TGGGGCGAGCTGGACCGCGGAGCGGAGCGCGCGCGGAGCCCGAGCAACGCGCG 798
|||||

Db 308 TGGGGCGAGCTGGACCGCGGAGCGGAGCGCGCGGAGCCCGAGCAACGCGCG 367
|||||

QY 799 CCCCCGAGCTGCGGAGCTGCGGCTTCGCGCGGAGGCTCGGCTCCCGAGGCGGGCC 858
|||||

Db 368 CCCCCGAGCTGCGGAGCTGCGGCTTCGCGCGGAGGCTCGGCTCCCGAGGCGGGCC 427
|||||

QY 859 CTGCTGGTGGTATTCACAGATCCCAAGCAAGAACCTGTTCCAGAGATGCGCGAGCAG 918
|||||

Db 428 CTGCTGGTGGTATTCACAGATCCCAAGCAAGAACCTGTTCCAGAGATGCGCGAGCAG 487
|||||

QY 919 CTGGGCTCGGCGAGGCTCGGCGCGCGGCGGCGGAGGCTGCTGGCGCGCGCG 978
|||||

Db 488 CTGGGCTCGGCGAGGCTCGGCGCGCGGCGGCGGAGGCTGCTGGCGCGCGCG 547
|||||

QY 979 TCGGGCGCGCGGATGCCAGGCTTGGCTGCGCTTCGCGCGCGCGCGCGCGCGCACG 1038
|||||

Db 548 TCGGGCGCGCGGATGCCAGGCTTGGCTGCGCTTCGCGCGCGCGCGCGCGCGCACG 607
|||||

QY 1039 GCTTCGCCAGTCGCCATGGCAAGCGGCGACGGCAAGAGTCCAGGCTACGCTGCAGCAAG 1098
|||||

Db 608 GCTTCGCCAGTCGCCATGGCAAGCGGCGACGGCAAGAGTCCAGGCTACGCTGCAGCAAG 667
|||||

QY 1099 AAGCCCTCGCAGTGAACCTCAAGGAGCTGGGCTGGGAGGAGCTGATTCGCGCCCTG 1158
|||||

Db 668 AAGCCCTCGCAGTGAACCTCAAGGAGCTGGGCTGGGAGGAGCTGATTCGCGCCCTG 727
|||||

QY 1159 GAGTACGAGGCTATCACGCGAGGCTGATGCGACTTCGCGCTCGCACCTGGAG 1218
|||||

Db 728 GAGTACGAGGCTATCACGCGAGGCTGATGCGACTTCGCGCTCGCACCTGGAG 787
|||||

QY 1219 CCACCAACACCGCATATCCAGAGCTGATGAAGTCCATCCATGACCCCGGCTCCACCCCG 1278
|||||

Db 788 CCACCAACACCGCATATCCAGAGCTGATGAAGTCCATCCATGACCCCGGCTCCACCCCG 847
|||||

QY 1279 CCAGCTGCTGGTCCCAACCAATGACTCCCATCAGATTCATACATCAGCGGGC 1338
|||||

Db 848 CCAGCTGCTGGTCCCAACCAATGACTCCCATCAGATTCATACATCAGCGGGC 907
|||||

QY 1339 AATAATGTTGTTACAAGCAGTACGAGGACATGGTGGTGGAGTCTGCGGCTGCAGGTAG 1398
|||||

Db 908 AATAATGTTGTTACAAGCAGTACGAGGACATGGTGGTGGAGTCTGCGGCTGCAGGTAG 967
|||||

QY 1399 CGGTGCTTTCCGCGCGCTTGGCGCG 1425
|||||

Db 968 CGGTGCTTTCCGCGCGCTTGGCGCG 994

RESULT 2
US-08-333-576C-25
; Sequence 25, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,576C
; FILING DATE: No. 6027919ember 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1171 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Human VL-1 protein
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..964
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 605..964

US-08-333-576C-25
Query Match 69.3%; Score 987; DB 3; Length 1171;
Best Local Similarity 100.0%; Pred. No. 2.7e-185;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCTCATGCTCTCA 498
|||||
Db 8 GATCTCTCGCACACTCCTCTCCGGAGACAGAAGTATTTGTTGATGTGCTCATGCTCTCA 67
|||||

QY 499 GACAAAGAAGAGCTGTTGGCGCGGAGCTGCGGCTCTTTTCGCCAGGCGCCCTCAGGCGCC 558
|||||

Db 68 GACAAAGAAGAGCTGTTGGCGCGGAGCTGCGGCTCTTTTCGCCAGGCGCCCTCAGGCGCC 127
|||||

QY 559 TGGGGGCCACACCGCGCGCTCCACGTGACGCTCTTCCTTTGCTTTCGCCCTTACTG 618
|||||

Db 128 TGGGGGCCACACCGCGCGCTCCACGTGACGCTCTTCCTTTGCTTTCGCCCTTACTG 187
|||||

QY 619 CTGGACGCGGACCTTGACCGCGAGGGCGCGCGCGGCTGGGAAGTCTTCGAC 678
|||||

Qy 1219 CCCACCAACACGCCATCATCCAGACGCTGATGAACCTCCATGGACCCCGCTCCACCCCG 1278
Dy 788 CCCACCAACACGCCATCATCCAGACGCTGATGAACCTCCATGGACCCCGCTCCACCCCG 847
Qy 1279 CCCAGCTGCTGCGTCCGCCCAAAATGACTCCCATAGCATTTATATACATCGAGCGGGC 1338
Dy 848 CCCAGCTGCTGCGTCCGCCCAAAATGACTCCCATAGCATTTATATACATCGAGCGGGC 907
Qy 1339 AATAATGTGTCTACAAAGCAGTACGAGACATGGTGTGAGTCTGCGGCTGCAGGTAG 1398
Dy 908 AATAATGTGTCTACAAAGCAGTACGAGACATGGTGTGAGTCTGCGGCTGCAGGTAG 967
Qy 1399 CGGTGCTTTCCCGCGCCCTTGCGCCG 1425
Dy 968 CGGTGCTTTCCCGCGCCCTTGCGCCG 994

RESULT 5

US-08-362-670B-31
; Sequence 31, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: MURINE MV2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..790
US-08-362-670B-31

Query Match 40.8%; Score 581.8; DB 1; Length 1046;
Best Local Similarity 87.2%; Pred. No. 7.4e-106;
Matches 675; Conservative 1; Mismatches 84; Indels 14; Gaps 3;

Qy 653 CGCGCGCGGTGGGAAGTC-TTCGACGTGTGGCAGGGCCTGCGCCACACGCCCTTGAAG 711
Dy 59 CCCAGCGCGGTGGGAAGTC-TTCGACGTGTGGCAGGGCCTGCGCCCTCAGCCTTGAAG 118
Qy 712 CAGTGTGCTTGGAGCTGCGGGCGCATGGGGCGAGCTGGACGCCGGGAGCGCGAGCG 771
Dy 119 CAGTGTGCTTGGAGCTGCGGGCGAGCCTGGGGTGGAGTCRCRCGGGATACCGGGCG 178
Qy 772 CGCGCGGGGACCCACAGCAACCGCCCGGACCTGCGGAGCTGGGCTTCGGCGCG 831
Dy 179 CGCGCGGGGTCCTCCACAGCACCCCGCTCTGACCTGCGGAGCTGGGCTTCGGTCGG 238
Qy 832 AGGTGCGGCTTCCCGAGGCGGCGCTGCTGTGTGTATTACACAGATCCACGCGAAG 891
Dy 239 AGGTGAGACCGCGCGAGGCGCGCTGCTGTGTGTATTACACAGATCCACGCGAAG 298
Qy 892 AACCTGTTCCGAGAGATGGCGGAGCAGCTGGGCTGCGGCGAGGCTGCGGGCGCG 951
Dy 299 AACCTGTTCACTGAGATGCATGAGCAGCTGGGCTCTCGAGAGGCT-----GCG 346
Qy 952 GCGCGCGAGGGTCTGCGCGCGCGCTGCGGGCGCGCGGATGCCAGGCTTGGCTGCC 1011
Dy 347 GGAGCGAGGGTCTGTCAGCGCGCTGCGGGCTCCCGACAGACCGGGCTCTTGGCTGCC 406
Qy 1012 TCGCCCGCGCGCGCGCGCGCGCTTCCGCGAGTCCCATGGCAAGCGCGCAGCG 1071
Dy 407 TCGCCCGCGCGCGCGCGCGCGCTTCCGCGAGGCTCAGCGCAAGCGACATGGC 466
Qy 1072 AAGAAGTCCAGGCTACGCTGCAGCAAGAGCCCTGCACGTGAACCTCAAGGAGCTGGC 1131
Dy 467 AAGAAGTCCAGGCTGCGCTGCAGCAAGAGCCCTGCACGTGAACCTTAAGGAGCTAGC 526
Qy 1132 TGGGACGACTGGATTATCGCGCGCTGAGTACAGGCTTACCTGCGAGGGTGTATGC 1191
Dy 527 TGGGACGACTGGATTATCGCGCGCTTACAGTACAGGCTTACCTGCGAGGGGTGTGC 586
Qy 1192 GACTTCCCGCTGCGCTCGCACCTGGAGCGCCACCAACCGCATCATCCAGACGCTGATG 1251
Dy 587 GACTTCCCGCTGCGCTCGCACCTGGAGCGCCACCAACCGCATCATCCAGACGCTGATG 646
Qy 1252 AACTCCATGAGCGCGCTCCACCGCGCGAGCTGCGTGGCGCCCAAAATGACTGCC 1311
Dy 647 AACTCCATGAGCGCGCTCCACCGCGCGAGCTGCGTGGCGCCCAAAATGACTGCC 706
Qy 1312 ATCAGCATTTATACATCGAGCGGCAATAATGTGTCTACAGCAGTAGAGGACATG 1371
Dy 707 ATCAGCATTTATACATCGAGCGGCAATAATGTGTCTACAGCAGTAGAGGACATG 766
Qy 1372 GTGTGGAGTCTGCGGCTGCGAGGTAGCGGTGCTTTCCCGCGCGCTTGGCGCG 1425
Dy 767 GTGTGGAGTCTGCGGCTGCGAGGTAGCGGTGCTTTCCCGCGCGCTTGGCGCG 819

RESULT 6

US-08-333-576C-31
; Sequence 31, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA

Qy	653	CGCGCGCGGCTGGGAAGTC - TTGCAGCTGTGGCAGGCGCTGCGCCACACAGCCCTTGAAG	711
Db	59	CGCGCGCGGCTGGGAAGTCTTTTCAGCTGTGGCAGGCGCTGCGCCCTTGAAG	118
Qy	712	CAGCTGTGTTGAGCTGCGGCGCCATGGGCGAGCTGGAGCGCGGGGAGCGCAGCGC	771
Db	119	CAGCTGTGCTTGAGTTGCGGCGAGCGCTGGGTGAGCTGGACRCGCGGGATACCGGGGCGC	178
Qy	772	CGCGCGGGGACCCACGAACCGCGCCCGCGACCTGCGGAGTCTGGGTCTGCGCGCGG	831
Db	179	CGCGGAGGGTCCCCAGCAGCACCGCCTCTGGACCTGCGGAGTCTGGGTCTCGTTCG	238
Qy	832	AGGTTGGGGCTCCCCAGGAGGGGCCCTGTGTGTGTTATTCACAGATCCACAGCGCAAG	891
Db	239	AGGTTGAGACGCCCGAGGAGCGGCCCTGCTGTGTAGTGTTCACAGATCGCAGCGCAAG	298
Qy	892	AACCTGTTCCAGAGATGCGCAGCAGCTGGGCTCGGCGGAGGCTGCGGGCCGGGCGCGC	951
Db	299	AACCTGTTCACTGAGATGATGAGCAGCTGGGCTCTGCAGAGGT-----GCG	346
Qy	952	GGCGCGAGGGTGTGTGGCGCGCGCGTGGGCGCGCCCGGATGCCAGGCTTGGCTGCC	1011
Db	347	GGAGCGAGGGTCAATGTCACGCGCGTGGGCTCCCGACAGACCGGCTTGGCTGCC	406
Qy	1012	TCGCGCGCGCGCGGGCGGCGCAGGCGCTTTCGCCAGTGCGCATGGAAGCGCACGCGC	1071
Db	407	TCGCGCGCGCGCGGGGAGCACCGCGCTTTCGCCAGCGCTCACGGCAAGCGACATGGC	466
Qy	1072	AAGAATCCAGGCTACGCTGCAGCAGAGAGCCCTGCAGCTGAACCTTCAGGAGCTGGC	1131
Db	467	AAGAATCCAGGCTGCGCTGCAGCAAGAAAGCCTTCGACGTGAATTTAAGGAGTTAGGC	526
Qy	1132	TGGGACACTTGGATTATCGCGCCCTTGGAGTACGAGGCGCTATCACTGCGAGGGTGTATGC	1191
Db	527	TGGGAGACTTGGATTATCGCGCCCTTAGAGTACGAGGCGCTATCACTGCGAGGGCTGTGC	586
Qy	1192	GACTTCCGCTGCGCTGCACTTGGAGCGCCACCAACCGCCATCATCCAGACGCTGATG	1251
Db	587	GACTTTCGCTGCGCTGCGACCTTGAGCGCAATAACCATGCATCATTCAGACGCTGATG	646
Qy	1252	AACTCATGACCCCGGCTCCACCCGCGCCAGCTGCTGGGTGCCACCAATTCACCTCC	1311
Db	647	AACTCATGACCCGGGCTCCACCCGCGCTGAGCTGCTGGGTGCCACCAATTCACCTCC	706
Qy	1312	ATCAGACTTCATACATCGACCGGGCAATAATGTGGTCTACAAGCAGTACGAGGACATG	1371
Db	707	ATTAGCATCTGTACATCGACCGGGCAATAATGTGCTACAAGCAGTATGAGGACATG	766
Qy	1372	GTGGTGGAGTGTGGCGCTGCAGGTAGCGGTGCCCTTTCGCGCGCGCTTGGCGCG	1425
Db	767	GTGGTGGAGTCTTGGCGCTGTAGGTAGCGGTG - CTGTCCCGCCACTGTGGCGAG	819

RESULT 8

```

PC/US94-14030A-31
? Sequence 31, Application PC/TUS9414030A
? GENERAL INFORMATION:
?
? APPLICANT: GENETICS INSTITUTE, INC.
? APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
? TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
?
? NUMBER OF SEQUENCES: 35
?
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: GENETICS INSTITUTE, INC.
? STREET: 87 Cambridgepark Drive
? CITY: Cambridge
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02140
?
? COMPUTER READABLE FORM:
?
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25

```



```

; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: HUMAN V1-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1301
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 990..1301
US-08-808-324-33

Query Match      18.2%; Score 259.2; DB 4; Length 1345;
Best Local Similarity 57.8%; Pred. No. 1.4e-42;
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;

QY 464 GACAGAAATATTGTTGATGCTCCATCTCTACACAAAGAGCTGGTGGCGCGG 523
Db 364 GCAGAGCTTCCTGTTGACGCTGTCAGCTTAACACACAGACAGAGTGGTGGTGC 423

QY 524 AGCTGCGGCTCTTTCCAGGCGCCCTCAGCGCCCTGGGGGCCACACGCGGCGCTCC 583
Db 424 AGCTGCGGCTGTCGCGCGGGATCTCCAGATC--GGCCCCAGGAGCTGACTTCTCC 481

QY 584 AGCTGAGCTCTTCCCTTTCCTTTGCCCCCTACTGCTGACGCGGAGCCCTGGACCCGC 643
Db 482 -----GCCGTTGCTGCTGCTGCTCACGTGCGCGCGCGCGCGAGCGCCACGCTGC 534

QY 644 AGGGGCGCGCGCGCGCTGGGAGTCTTCGACGTGTGGCAGGCGCTGCCACCACC 703
Db 535 TGACTCGCGGCGAGCTGAGCCCTAGTCCGCTGAGGCTGGGAGGCTTCGACGTGGGG 594

QY 704 CTGGAAGCAGCTGTGTTGGAGCTGCGGCGCGCATGGGCGGAGCTGGACGCGGGGAG 763
Db 595 ACGCCATGAGCGCCACCGCTGTAACCGCGCCCCCGCGGCTTCTGCCTTGTGCTC 654

QY 764 CGAGGCGCGCGGGGAGACCCAGAACCGCGCGCCCCCGGACCTTGGAGTCTGGGCT 823
Db 655 GCGCATGTGCGAGCGCGGCTGCGGAGCGCCGTTGGCACCTGCGGCGACTTGGGCTGCG 714

QY 824 TCGCGGAGGCTGGGCGCTCCCGAGGAGCGGCGCTGCTGCTGATTCACACAGATCC 883
Db 715 CGGGCGGAGGGGCTGTCGGGAGAGAGGAGCGCGGCTGCTGCTGCTCTCCCGCACGC 774

QY 884 AGCGCA---AGAACCTGTTGCGAGATGCGGAGCAGCTGGGCTCGGCGGAGGCTGC-- 938
Db 775 AGAGGAAGAGAGCTTATTCCGAGATCCGCGCCAGCGCGCGCTCGGGCGCGCTC 834

QY 939 -----GGGCGCGCGCGCGCGGAGGCTGCTGCTGCGCGCGCGCTCGGGCGCGCC 988
Db 835 TGGCCTCAGAGCGCTGCCCGACCCAGGAACCGGACCGCTCGGCAAGGGCAGTCA 894

QY 989 CGATGCCAGGCTTGGCTGCCCTCCCGCGCGCGCGCGCGCGCGCGCTTCGCCA 1048
Db 895 GCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 954

QY 1049 GTCGCATGCGAAGCGGCGACGCAAGATGCTCAGGCTACGCTGCGAGCAAGAGCCCTGC 1108
Db 955 GGGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1014

QY 1109 AGCTGAATTCAGAGCTGGGCTGGGAGCTGATTCGCGCCCTTGGAGTACGAGG 1168
Db 1015 AGCTGAGCTTCAAGAGCTGCGCTGGGAGCTGATGATCATCGCGCGCTGCTACGAGG 1074

QY 1169 CTTATCACTGCGAGGCTGATCGACTTCCGCTCGGCTCGGCTCGGAGCCCAACC 1228
Db 1075 CGTACACTGCGAGGCGCTTTCGACTTCTCCCTTTCGCTTTCGACCTTCGAGCCCAACC 1134

QY 1229 AGCCCATCATCAGAGCTGATGAATCATGAGACCGGGCTTCCACCCCGCGCGAGCTGCT 1288
Db 1135 ATGCCATCATTCAGAGCTGTCTCAACTCCATGCGCACAGACGCGCGCGCGCTTCTGCT 1194

QY 1289 GCGTCCCAACCAATGACTTCCATCAGATTCATATCATGACGCGGGGCAATATGTGG 1348
Db 1289 GCGTCCCAACCAATGACTTCCATCAGATTCATATCATGACGCGGGGCAATATGTGG 1348
```

```

Db 1195 GTGTGCCAGCGCGCTCAGCCCCATCAGCATCTCTACATCAGCGCGCCAAACAGTTG 1254
QY 1349 TCTACAAGCAGTACGAGGACATGGTGGTACTGCTGGCGCTGCAGGTAGCG 1400
Db 1255 TCTACAAGCAATACGAGGACATGGTGGTGGCGCTTGGCGCTGCAGGTAGCG 1306

RESULT 15
PCT-US94-14030A-33
; Sequence 33, Application PC/TUS9414030A
; GENERAL INFORMATION:
; APPLICANT: GENETICS INSTITUTE, INC.
; APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLEGE
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14030A
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/164,103
; FILING DATE: 07-DEC-1993
; APPLICATION NUMBER: US 08/217,780
; FILING DATE: 25-MAR-1994
; APPLICATION NUMBER: US 08/333,576
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202D-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1345 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; CLONE: HUMAN V1-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 138..1301
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 990..1301
PCT-US94-14030A-33
```

```

Query Match      18.2%; Score 259.2; DB 5; Length 1345;
Best Local Similarity 57.8%; Pred. No. 1.4e-42;
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;

QY 464 GACAGAAATATTGTTGATGCTCCATGCTCTACAGAAAGAGCTGGTGGCGCGG 523
Db 364 GCAGAGCTTCTGTTTCACGCTTCCAGCTTAACACACAGACAGAGTGGTGGGTCGG 423

QY 524 AGCTGCGGCTCTTTTGGCCAGGCGCCCTCAGGCGCCCTCGAGGCGCGCTTGGGGCGCGCTCC 583
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 01:02:45 ; Search time 86 Seconds
(without alignments)
6275.472 Million cell updates/sec

Title: US-09-825-751A-19
Perfect score: 1425
Sequence: 1 cctctggggagacgcagcca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues
Total number of hits satisfying chosen parameters: 6711156

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987	69.3	1171	9	US-09-945-182-25
2	928.2	65.1	1308	10	US-09-730-772-12
3	928.2	65.1	1308	10	US-09-735-849-12
4	581.8	40.8	1046	9	US-09-945-182-31
5	259.2	18.2	1345	9	US-09-945-182-33
6	253	17.8	2703	8	US-08-981-490B-2
7	251.6	17.7	1207	9	US-09-945-182-3
8	251.4	17.6	2341	10	US-09-730-772-11
9	251.4	17.6	2341	10	US-09-735-849-11
10	245.8	17.2	357	8	US-08-945-459A-4
11	245.8	17.2	357	9	US-09-068-253-1
12	245	17.2	1203	9	US-09-945-182-29
13	244.8	17.2	926	9	US-09-945-182-1
14	239	16.8	2329	9	US-09-880-708-9
15	237	16.6	475	10	US-09-833-381-582
16	232	16.3	272	9	US-09-945-182-12
17	230	16.1	1233	9	US-09-945-182-27
18	208.8	14.7	3585	9	US-09-945-182-9
19	195.2	13.7	272	9	US-09-945-182-10

20	172.8	12.1	272	9	US-09-945-182-14
21	144	10.1	1346	10	US-09-784-911-11
22	144	10.1	1352	10	US-09-784-911-13
23	138.4	9.7	557	10	US-09-765-527-248
24	138.4	9.7	557	10	US-09-760-397-1
25	138.4	9.7	1822	8	US-08-822-186-1
26	138.4	9.7	1822	8	US-08-937-755-1
27	138.4	9.7	1822	9	US-09-982-543A-9
28	138.4	9.7	1822	10	US-09-045-331-1
29	138.4	9.7	1822	10	US-09-828-607-1
30	138.4	9.7	1878	10	US-09-954-456-1824
31	136.2	9.6	610	10	US-09-760-397-7
32	136.2	9.6	613	10	US-09-765-527-256
33	136.2	9.6	613	10	US-09-765-527-260
34	136.2	9.6	646	10	US-09-760-397-13
35	136.2	9.6	658	10	US-09-765-527-254
36	136.2	9.6	661	10	US-09-765-527-262
37	131	9.2	1350	10	US-09-784-911-7
38	131	9.2	1362	10	US-09-784-911-3
39	130.8	9.2	1353	10	US-09-784-911-9
40	114.8	8.1	1926	12	US-10-044-090-28
41	114.4	8.0	1224	10	US-09-749-728B-70
42	111	7.8	1547	9	US-10-044-716-1
43	111	7.8	14759	10	US-09-952-360-1
44	107.2	7.5	1289	10	US-09-784-911-5
45	106.8	7.5	2003	10	US-09-764-869-1535

ALIGNMENTS

RESULT 1
US-09-945-182-25
; Sequence 25, Application US/09945182
; Patent No. US20020160494A1
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; Wozney, John
; Rosen, Vicki A.
; Wolfman, Neil
; Thomsen, Gerald H.
; Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/945,182
; Filing DATE: 31-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,324
; Filing DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1171 base pairs

Sequence 14, Appl
Sequence 11, Appl
Sequence 13, Appl
Sequence 248, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1824, Ap
Sequence 7, Appl
Sequence 256, App
Sequence 260, App
Sequence 13, Appl
Sequence 254, App
Sequence 262, App
Sequence 7, Appl
Sequence 3, Appl
Sequence 9, Appl
Sequence 28, Appl
Sequence 70, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 5, Appl
Sequence 1535, Ap

```
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ IMMEDIATE SOURCE:
/ CLONE: Human VL-1 protein
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 2..964
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 605..964
/ SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-945-182-25

Query Match      69.3%; Score 987; DB 9; Length 1171;
Best Local Similarity 100.0%; Pred. No. 5.7e-192;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAAATTTGTTGATGTGTCATGCTCTCA 498
      |||
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAAATTTGTTGATGTGTCATGCTCTCA 67

QY 499 GACAAAGAGAGCTGTGGGCGGAGCTGCGGCTCTTTGCGCAGGCGCCTCAGCGCCC 558
      |||
Db 68 GACAAAGAGAGCTGTGGGCGGAGCTGCGGCTCTTTGCGCAGGCGCCTCAGCGCCC 127

QY 559 TGGGGCCACACAGCGCGGCTCCACGTGACGCTCTCCCTTGCCCTTCGCGCCCTACTG 618
      |||
Db 128 TGGGGCCACACAGCGCGGCTCCACGTGACGCTCTCCCTTGCCCTTCGCGCCCTACTG 187

QY 619 CTGGAGCGCGGACCTTGAACCGGAGGCGCGCGCGCGGCTGGGAAGTCTTCGAC 678
      |||
Db 188 CTGGAGCGCGGACCTTGAACCGGAGGCGCGCGCGCGGCTGGGAAGTCTTCGAC 247

QY 679 GTGTGGAGGGCTGTGCCACACAGCCCTTGAAGACAGCTGTGCTTGAGCTGCGGGCCGA 738
      |||
Db 248 GTGTGGAGGGCTGTGCCACACAGCCCTTGAAGACAGCTGTGCTTGAGCTGCGGGCCGA 307

QY 739 TGGGGGAGCTGACCGCGGAGGCGGAGCGCGCGCGGAGCCCGAGCAACCGCGG 798
      |||
Db 308 TGGGGGAGCTGACCGCGGAGGCGGAGCGCGCGCGGAGCCCGAGCAACCGCGG 367

QY 799 CCCCCGACCTGCGGAGCTTGGGCTTTCGCGGAGGCTTCGCGGCTCCCGAGGCGGGCC 858
      |||
Db 368 CCCCCGACCTGCGGAGCTTGGGCTTTCGCGGAGGCTTCGCGGCTCCCGAGGCGGGCC 427

QY 859 CTGCTGGTGTATTACAGATCCAGCGGAGAACTCTGTTCGACAGATGCGCGAGCAG 918
      |||
Db 428 CTGCTGGTGTATTACAGATCCAGCGGAGAACTCTGTTCGACAGATGCGCGAGCAG 487

QY 919 CTGGGCTCGCGGAGCTGCGGCGCGGCGGCGGCGGCGGAGGCTGCGCGCGCGCG 978
      |||
Db 488 CTGGGCTCGCGGAGCTGCGGCGCGGCGGCGGCGGCGGAGGCTGCGCGCGCGCG 547

QY 979 TCGGGCGCGCGGATCCAGGCTTGGCTTGCCTTCGCGGCGCGCGGCGGCGGCGCAG 1038
      |||
Db 548 TCGGGCGCGCGGATCCAGGCTTGGCTTGCCTTCGCGGCGCGCGGCGGCGGCGCAG 607

QY 1039 GCGTTGCGAGTGGCCATGGCAAGCGGACGCGGCAAGAGTCCAGGCTACGCTGCAGCAAG 1098
      |||
Db 608 GCGTTGCGAGTGGCCATGGCAAGCGGACGCGGCAAGAGTCCAGGCTACGCTGCAGCAAG 667

QY 1099 AAGCCCTCGACGTGACCTCAAGGAGCTGGGCTGGGAGCTGGATTATCGGCCCCCTG 1158
      |||
Db 668 AAGCCCTCGACGTGACCTCAAGGAGCTGGGCTGGGAGCTGGATTATCGGCCCCCTG 727

QY 1159 GAGTACGAGGCTATCAGTCGAGGCTGTATGCGACTTCCCGCTGCGCTCGCACCTGGAG 1218
      |||
Db 728 GAGTACGAGGCTATCAGTCGAGGCTGTATGCGACTTCCCGCTGCGCTCGCACCTGGAG 787

QY 1219 CCGACCAACCGCCATCATCCAGAGCTGTGAACTCCATGACCGCGGCTCCACCGCG 1278
      |||
```

```
Db 788 CCACCAACACAGCCATCCAGACGCTGATGAATCCATGACCCCGGCTCCACCCG 847
QY 1279 CCAGCTGCTGCTGCCACCAAAATTTGACTCCCATCAGCAATTTATACATCGACGGGC 1338
      |||
Db 848 CCAGCTGCTGCTGCCACCAAAATTTGACTCCCATCAGCAATTTATACATCGACGGGC 907
      |||
QY 1339 AATAATGTGCTTACAAAGCAGTACGAGGACATGTTGGTGAGTCTGCGGCTCAGGTAG 1398
      |||
Db 908 AATAATGTGCTTACAAAGCAGTACGAGGACATGTTGGTGAGTCTGCGGCTCAGGTAG 967
      |||
QY 1399 CGGTGCTTTCCGCGCGCTTGGCCG 1425
      |||
Db 968 CGGTGCTTTCCGCGCGCTTGGCCG 994
      |||

RESULT 2
US-09-730-772-12
; Sequence 12, Application US/09730772
; Patent No. US20010011131A1
; GENERAL INFORMATION:
; APPLICANT: Luyten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; APPLICANT: Chang, Steven Chao-Huan
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/730,772
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,081
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartfield, Neil S
; REGISTRATION NUMBER: 39,901
; REFERENCE/DOCKET NUMBER: NIH099,001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1308 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-730-772-12

Query Match      65.1%; Score 928.2; DB 10; Length 1308;
Best Local Similarity 86.5%; Pred. No. 5e-180;
Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;

QY 128 CCGTCCGCGGCTGGGTTCCACCAAGGGCATCGAAGCGCCCAAGGAAGCAAGATCGACG 187
      |||
Db 5 CCGTCCGCGGCTGGGTTCCCGCAAGGGATCGGAACGCGCAAGGAAGGAAGATCGCCG 64

QY 188 GGGCGCGCGCGACAGTACGCGGGCGCGGAG-----GGCCAGGAACACAGCCG 238
      |||
Db 65 GGGCGCGGAGAGAGATGCCAGCGCGCGGAGCCCTGATGCCAGGAGCCCCCGCGA 124
```

QY 239 GGCTTCAGGACGAAACCC-----CGGGCTCAGCAGCCCGGGGCGAGAGCCGCCAG 289
Db 125 GGCGCAGGAGGAGCCCGGCGGCGCCACAGCAGCCTGAAGCTCGGAGCCTCCCG 184
QY 290 GCAGGGTCCGGCGTGGTGGCCCGCAGAGTACATGTCTCAATCTACAGGACTTACTCCA 349
Db 185 GCAGGGCGCGCGCTTGGTGGCCCGCAGAGTACATGTCTCAATCTACAGGACTTACTCCA 244
QY 350 TCAGTGAAGAGCTGGGATCAATGTCAGCTTTTCCAGTCTTCCAGTCCGGCTAATACGA 409
Db 245 TCAGGAGAGAGCTGGGATCAATGCTAGCTTTTCCAGTCTTCCAGTCCGGCTAATACGA 304
QY 410 TCACAGCTTTGTAGCAGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGA 469
Db 305 TCAGTGTGTAGCAGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGA 364
QY 470 AGTATTTGTGTGTGTCATGCTCTCAGACAAAGAGCTGGTGGGCGGAGGTGC 529
Db 365 AGTATTTGTGTGTGTCATGCTCTCAGACAAAGAGCTGGTGGGCGGAGGTGC 424
QY 530 GGCTCTTTCAGGAGGCGCTCAGCGCCCTGGGGGCGCACAGCGGCGGCTCCAGCTGC 599
Db 425 GGCTGTTCAGGAGGCGCGCTGCGCTGGCGCGCGCGC-----GGCGCTCCGCTTCG 479
QY 590 AGCTCTTCCCTTTCGCTTTCGCTGACGCGGAGCCCTGGAGCCCGCAGGGG 649
Db 480 AGCTCTTTCGCTGCCAGTCCGCTGCTGGAAGCGGAGCCTGGACCCGCA-GGG 538
QY 650 CGCGCGCGCGGCTGGAGTGTTCAGCTGTGGAGGCGCTGCGCCACAGCCCTGGA 709
Db 539 CGCGCGCGCGGCTGGAGTGTTCAGCTGTGGGCGGCTGCGCCCGCCCGCTGGA 598
QY 710 AGCAGCTGTGTGGAGCTCGCGGCGCATGG---GSCGAGCTGGAGCGCGGGAGGCGG 766
Db 599 AGCAGCTGTGTGGAGCTTCGGGCGCGCTGGGCGGAGCGGCGCGGAGGAGCG 658
QY 767 AGCGCGCGCGGAGCCCGCAGAACCGCGCGCGCGGACCTGCGGAGTCTGGGCTTCG 826
Db 659 AGCGCGCAGCGCTGGGCGCGCAGCAGCGCGCGCGGACCTGCGGAGTCTGGGCTTCG 718
QY 827 CGCGGAGGTGCGGCTTCCCGAGAGGGGCGCTGCTGTGTGTATTCACAGATCCGAGC 886
Db 719 GCGGAGGTGCGGAGCGCGCGCGCGGCGGCTGCTGTGTGTATTCACAGATCCGAGC 778
QY 887 GCAAGAACCTGTTTCGAGAGATCGCGAGCAGCTGGGCTCGG---CGAGGCTCGGGCC 943
Db 779 GCAAGAACCTGTTTCGCGAGATCGCGAGCAGCTGGGCTCGGCGACCGAGGTGTCGGCC 838
QY 944 CGGCGCGGCGCGGAGGGTCTGTG-----GCCCGCGCGCTCGG 982
Db 839 CGGCTGTGGGCGGAGGGTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 898
QY 983 CGCGCGGAGTGCAGGCTTGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1042
Db 899 GCACCGCGGAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
QY 1043 TCAGCAGTCCCATGCAAGCGGCGAGGAGTCCAGGCTACGCTGACGAGCAAGAGC 1102
Db 956 TCAGCAGTCCCATGCAAGCGGCGAGGAGTTCAGGCTGCGCTGCGAGCAAGAGC 1015
QY 1103 CCTGCACTGAATTCAGGAGCTGGGCTGGGAGCAGTGGATATTCGCGCGCGCTGGAGT 1162
Db 1016 CCTGCACTGAATTCAGGAGCTGGGCTGGGAGCAGTGGATATTCGCGCGCGCTGGAGT 1075
QY 1163 ACAGGCTTATCTAGGAGGTGTATGGAGCTTCCGCTGCGCTGCGAGTCCGAGCCCA 1222
Db 1076 ACAGGCTTATCTAGGAGGTGTATGGAGCTTCCGCTGCGCTGCGAGTCCGAGCCCA 1135
QY 1223 CCAACAGCGCATTCAGAGCTGATGAACCTCCATGACCCCGGCTCCACCCCGCCCA 1282
Db 1136 CCAACAGCGCATTCAGAGCTGATGAACCTCCATGACCCCGGCTCCACCCCGCCCA 1195

QY 1283 GCTGCTCGTGGCCCGCCACCAATGACTCCCATCAGCATCTTATATCATGACGCGGCAATA 1342
Db 1196 GCTGCTCGTGGCCCGCCACCAATGACTCCCATCAGCATCTTATATCATGACGCGGCAATA 1255
QY 1343 ATGTGCTCTACAAGCAGTACGAGGACATGTGTGGAGTCTGCGGCTGCAGG 1395
Db 1256 ATGTGCTCTACAAGCAGTACGAGGACATGTGTGGAGTCTGCGGCTGCAGG 1308
RESULT 3
US-09-735-849-12
; Sequence 12, Application US/09735849
; Patent No. US20010037017A1
; GENERAL INFORMATION:
; APPLICANT: Luyten, Frank P.
; APPLICANT: Moos, Jr., Malcolm
; TITLE OF INVENTION: CARTILAGE-DERIVED MORPHOGENETIC
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/735,849
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/836,081
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bartfeld, Neil S
; REGISTRATION NUMBER: 39,901
; REFERENCE/DOCKET NUMBER: NIH099,001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1308 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-735-849-12
Query Match 65.1%; Score 928.2; DB 10; Length 1308;
Best Local Similarity 86.5%; Pred. No. 5e-180;
Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;
QY 128 CGTCCCGCGAGCTGGGTTCACCAAGGCGATGCGAGCCCGCAAGGAGGCAAGATGCAGC 187
Db 5 CGTCCCGCGAGCTGGGTTCGCGCCAGGGAATGCAACGCCCAAGGAGGAGGATGCCGC 64
QY 188 GGGCGCGCGCGACAGTACGCGCGGGCGGGAG-----GGCCAGGAACACAGCCGC 238
Db 65 GGGCGCGCGAGAGAAATGCGACGCGCGGGAGCCCTGGATCGCAGGAGCCCGCGCGA 124
QY 239 GGCTTCAGGACGAAACC-----CGGGCTCAGCAGCCCGGGCGCAGAGCCCGCAG 289
Db 125 GGCGCGGAGGAGGAGCCCGAGCGGCGCGCACAGCCTGAAGCTCGGGAGCGCTCCCG 184
QY 290 GCAGGGGTCCGCGGTGGTGGCCCGCAGAGTACATGCTGTCAATCTACAGGACTTACTCCA 349


```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
;
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
;
US-09-735-849-11

Query Match 17.6%; Score 251.4; DB 10; Length 2341;
Best Local Similarity 78.7%; Pred. No. 1.2e-42;
Matches 300; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1021 CGCCGGCGGGCGGCGGCGCTTCGCCAGTTCGCCATGGCAAGCGGCACGCAAGAAAGTCC 1080
DB 1393 CGCGAAAGCGGGGCGCCCATCGGCCACTCGCCAGGCAAGCGCCAGCAAGACCTT 1452
QY 1081 AGCTACGCTGCAGCAAGAGCCCTTCAGGTGAACCTTCAAGAGAGCTGGCTGGGAGAC 1140
DB 1453 AAGGCTCGTGCAGTCGGAAGGACACTGCATGTCAACTTCAAGGACATGGGCTGGGAGAC 1512
QY 1141 TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGCGAGGGTGTATGCGACTTCCCG 1200
DB 1513 TGGATATCGCACCCCTTGATGACGAGGCTTCCACTGCGAGGGGCTGCGGAGTCCCA 1572
QY 1201 CTGCGCTCGACCTGGAGGCCACCAACCGCCATCATCCAGACGCTGATGAACCTCATG 1260
DB 1573 TTGCGCTCCACCTGGAGGCCACGAATATGCAATCATCAGACCTGATGAACCTGATG 1632
QY 1261 GACCCGGCTCCACCCGCGCCAGCTCTCGTGGTCCCAACCAATGACTCCCATCAGCAT 1320
DB 1633 GACCCGGAGTCCACACCCACCTCTGTGTGCGCCACGCGGTGAGTCCCATCAGCATC 1692
QY 1321 CTATACATCGACGCGGGCAATATGTGGTCTACAAGCAGTACGAGGACATGTTGGTGGAG 1380
DB 1693 CTCTTATTGACTCTCGCAACACGTTGGTGTATTAAGCAGTATGAGGACATGTTGCTGGAG 1752
QY 1381 TCGTGGGCTCGAGGTAGCGG 1401
DB 1753 TCGTGTGCTGCAGTAGCAG 1773

RESULT 10
US-08-945-459A-4
; Sequence 4, Application US/08945459A
; Patent No. US20020102633A1
; GENERAL INFORMATION:
; APPLICANT: MAKISHIMA, FUSAO; TAKAMATSU,
; APPLICANT: HIROYUKI; MIKI, MICHIO; KAWAI,
; APPLICANT: SHINJI; KIMURA, HIRO; MATSUMOTO,
; APPLICANT: TOMOKI; KATSUURA, MIEKO; ENOMOTO,
; APPLICANT: KOICHI; SATOH, YUSUKE
; TITLE OF INVENTION: A NOVEL PROTEIN AND
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS
; ADDRESS: LLP
; STREET: 600 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/945,459A
; FILING DATE: 09-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/01062
; FILING DATE: 19-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP7/322403
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP7/93664
; FILING DATE: 19-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146,1275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
;
US-08-945-459A-4

Query Match 17.2%; Score 245.8; DB 8; Length 357;
Best Local Similarity 81.0%; Pred. No. 1.2e-41;
Matches 286; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1043 TCGCCAGTCGCGATGCGCAAGCGCGCAGCAAGAGTCCAGCTACGCTCGCAGCAAGAAGC 1102
DB 5 TCGCCACTCGCGAGGCAAGCGACCCAGCAAGAACCTTAAGGCTCGCTCGAGTCGGAAGG 64
QY 1103 CCTCGACGTGAACCTTCAAGGAGCTGGCTGGGAGCACTGGATTATCGCGCCCCCTGGAGT 1162
DB 65 CACTGCATGTCAACTTCAAGGACATGGGCTGGGAGCACTGGATCATCGCACCCCTTGAGT 124
QY 1163 AGAGGCTTACATCGCGAGGGTGTATGCGACTTCCCGTGGCTGGCTGGAGCCCA 1222
DB 125 AGAGGCTTTCACACTGCGAGGGGCTGTGCGAGTTCCCATTTGCGCTCCCACTGGAGCCCA 184
QY 1223 CCAACACGCGCATCATCCAGAGCTGATGAACCTCACTGAGACCCCGGCTCCACCCGCCCA 1282
DB 185 CGATCATCGAGTCCAGACCCCTGATGAACCTCACTGAGACCCCGGAGTCCACACCCCA 244
QY 1283 GCTGCTGCTGCGCCCAAAATTGACTCCCATCAGCATTTATATACATCGAGCGGGCAATA 1342
DB 245 CCGTGTGTGTGCGCCAGCGACTGAGTCCCATCAGCATCTCTTTCATTGACTCTGCCAACA 304
QY 1343 ATGTGCTGTACAAGCAGTACGAGGACATGGTGGTGGAGTCGTGGCGCTGCAGG 1395
DB 305 ACGTGGTGTATAAGCAGTATGAGGACATGGTGTGGAGTCGTGGCTGCAGG 357

RESULT 11
US-09-068-253-1
; Sequence 1, Application US/09068253
; Patent No. US2002016831A1
; GENERAL INFORMATION:
; APPLICANT: SHIMURA, Takesada
; APPLICANT: TORIYAMA, Satsuki
; TITLE OF INVENTION: CARTILAGE/ BONE INDUCING MATERIALS FOR REPARATION
; FILE REFERENCE: 146,1286
; CURRENT APPLICATION NUMBER: US/09/068,253
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: PCT/JP96/03333
; PRIOR FILING DATE: 1996-11-14
; PRIOR APPLICATION NUMBER: JP 7/322402
; PRIOR FILING DATE: 1995-11-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
```



```
; SEQ ID NO 1
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(357)
; OTHER INFORMATION: Relevant amino acid residues in SEQ ID NO: 1 from
; PUBLIC INFORMATION: 1 to 119 in WO 95/04819
; PUBLICATION INFORMATION:
; AUTHORS: HOTTEN, Gertrud
; AUTHORS: NEIDHARDT, Helge
; AUTHORS: PAULISTA, Michael
; TITLE: NEW GROWTH/DIFFERENTIATION FACTOR OF THE TGF-BETA
; TYPE: FAMILY
; PATENT DOCUMENT NUMBER: WO 95/04819
; PATENT FILING DATE: 1995-02-16
; RELEVANT RESIDUES: 1 TO 119
US-09-068-253-1

Query Match 17.2%; Score 245.8; DB 9; Length 357;
Best Local Similarity 81.0%; Pred. No. 1.2e-41;
Matches 286; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 1043 TCCCGATGCGCATGCAAGCGCCGCAAGAGTCCAGGCTACGCTGCAGCAAGAAGC 1102
DB 5 TGGCCACTCGCCAGGGGCAAGCACCAGCAAGAACCTTAAGGCTCGCTGCAGTCGGAAG 64

QY 1103 CCTGCAGCTGAACCTCAGGAGCTGGGTGGGACGACTGGATTATATCGGCCCTGGAGT 1162
DB 65 CACTGCATGTCAACTTCAAGGACATGGGTGGGACGACTGGATTCATCGCACCCCTTGAGT 124

QY 1163 ACAGGCTTATCACTCGAGGGTGTATGCGACTTCCCGTGCCTCGCACCTGGAGCCCA 1222
DB 125 ACAGGCTTTCACACTCGAGGGGCTGTGCGAGTTCCATTCGCTCCACCTGGAGCCCA 184

QY 1223 CCAACACGCGCATATCCAGAGCTGATGAACCTCCTATGAGACCCCGCTCCACCCGCCCA 1282
DB 185 CGAATCATGCAATCCAGACGCTGATGAACCTCCTATGAGACCCCGCTCCACACCCCA 244

QY 1283 GCTGCTGCGTCCCAACATTCATCCCATCAGCATTCATACATCGAGCGGGCAATA 1342
DB 245 CTTGCTGTGTGCCCGCAGCGACTGAGTCCCATCAGCATCTCTTCATTCAGTCTGCAACA 304

QY 1343 ATGTGCTACAAAGCAGTACGAGACATGTTGTTGAGTCTGTCGGCTGTCAGG 1395
DB 305 ACCTGGTGTATAGCAGTATGAGGACATGTTGTTGAGTCTGTCGGCTGTCAGG 357

RESULT 12
US-09-945-182-29
; Sequence 29, Application US/09945182
; Patent No. US200201604941
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; Wozney, John
; Rosen, Vicki A.
; Wolfman, Neil
; Thomsen, Gerald H.
; Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/945,182
; FILING DATE: 31-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,324
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: murine MV1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..721
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-945-182-29

Query Match 17.2%; Score 245; DB 9; Length 1203;
Best Local Similarity 71.2%; Pred. No. 2.1e-41;
Matches 351; Conservative 0; Mismatches 137; Indels 5; Gaps 2;

QY 912 CGAGCAGCTGGGCTCGCGCAGGCTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 971
DB 233 CCAGGCGCTGGGTACGAAAGCAACCTGGGCGCTCGGAGCGCGAGGACTGCGCTG 292

QY 972 GC-CGCGGTGCGGCGCGCGCGGATGCCAGGCTTGGTGTGCGCGCGCGCGCGCGCGCG 1030
DB 293 GCTGGGACTCGGGGAGNAGGAAACGGTGTGGCGCGGTGGCGCGCGCGCGCGCGCGCG 352

QY 1031 GGC- ---GCACGCGCTTCGCCAGTCCCATGCGCAAGCGCACGCGCAAGAGTCCAGGCTA 1086
DB 353 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 412

QY 1087 CGCTGAGCAAGAAAGCCCTGCACGTGAACCTCAAGAGCTGGGCTGGGACGACTGGATT 1146
DB 413 CGCTGCGGTGCGAAAGTCACTGCACGTGGACTTTAAGGAGCTGGGCTGGGACGACTGATC 472

QY 1147 ATCGCGCGCGCTGGAGTACGAGGCTATCACTGGGAGGGTGTATGCGACTTCCCGCTGCGC 1206
DB 473 ATCGCGCGCTTAGACTACGAGGCTATACACTGGAGGGGTTTGCACCTTCTCTGCGCG 532

QY 1207 TCGCAGCTGGAGCGCCCAACCAACGCGCATATCCAGACGCTGATGAATCCCATGGACCCC 1266
DB 533 TCGCAGCTGGAGCGCTACCAACCAACGCGCATATTCAGACGCTGTCTCAACTCCATGGCGCC 592

QY 1267 GGCTCCACCGCGCGCGCGCTGCTGCGTGGCCCAACCAATTCAGTCCCATCAGCATTCATAC 1326
DB 593 GAGCGTGGCGCGCGCGCTTCTGCTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 652

QY 1327 ATCGAGCGCGGCAATAATGTGGTCTACAAGCAGTACGAGGACATGTTGGTGGAGTGGTGC 1386
DB 653 ATCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 712

QY 1387 GGCTGAGGTAGC 1399
DB 713 GGCTGAGGTAGC 725

RESULT 13
US-09-945-182-1
```

Qy	939	-----GGGCCCCGGGGCGGGCGCGAGGGGTCTGTGGCCGCGCCCGTCTGGGGCG	986
Db	414	TCTGGCTCAGAGCCGTCGCCGACCCAGGAACCGGCACCGCTGCCCAAGGGCAGTCAT	473
Qy	987	CCGGATGCCAGGCGTTGGCTCGCCCGCGCGCGGGCGGGCGGCGACAGGCTTCGC	1046
Db	474	TGGGGCCCGAGAGGAGAGACGGCTGTGGCCGGGACGGGACAGCGAGGGCAGGGG	533
Qy	1047	CAGTCGCCATGGCAAGCGGCACGGCAAGATCCAGGCTACGCTGCAGCAAGAACGCCCT	1106
Db	534	CGGGGGCGGGCGCGGCCACGGGCGCAGGGGCGGAGCGCTGCAGCGCAAGACCGGTT	593
Qy	1107	GCAGGTGAAGCTCAAGAGAGCTGGCTGGGACAGCTGGATTATTCGGCGCCCTTGGAGTACGA	1166
Db	594	GCAGGTGAAGCTCAAGAGAGCTCGGCTGGGACAGCTGGATCATCGCGCGCGTGGAGTACGA	653
Qy	1167	GGCTATACCTGCGAGGGTGTATGCGACTTCGCGCTCGCGCTCGACCTGGAGCCACACAA	1226
Db	654	GGGTACCACTGGAGGGCTTTGGCACTTCCTTTGGTTTCGCACCTCGAGCCACCA	713
Qy	1227	CCAGCGCATCATCCAGACGCTGATGAATCCATAGACCCCGGGCTCCACCCGCGCAGCTG	1286
Db	714	CCATGCCATCATTCAGACGCTGCTCAACTCCATGGCATCGACAGCGGGCGCGGCTCTCTG	773
Qy	1287	CTGGTGGCCCAACAAATTGACTCCCATCAGCATTCATATACATCGACGGGGCAATAATGT	1346
Db	774	CTGTGTGCCAGCGGGCTCAGGCCCATCAGCATTCCTCTACATCGACGGCGCCCAACAGT	833
Qy	1347	GGTCTAAGACGATCAGGAGCATGGTGGTGGAGTCGTGCGGCTGCAGGTAGCG	1400
Db	834	TGTCTACAAGCAATACGAGGACATGGTGGTGGAGGCTTCGCGCTGCAGGTAGCG	887

RESULT 14
US-09-880-708-9
; Sequence 9, Application US/09880708
; Patent No. US20020165361A1
; GENERAL INFORMATION:
; APPLICANT: Lee, Se-Jin
; Huynh, Thanh
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP
; STREET: 4365 Executive Drive, Suite 1600
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121-2189
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/880,708
; FILING DATE: 12-Jun-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/145,060
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/003,144
; FILING DATE: 12-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lisa A. Haile, Ph.D.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/057002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858/677-1456
; TELEFAX: 619/677-1465
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2329 base pairs
; TYPE: nucleic acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: GD-5
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 322...1806
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-880-708-9

Query Match      16.8%; Score 239; DB 9; Length 2329;
Best Local Similarity 74.9%; Pred. No. 3.8e-40;
Matches 299; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1021 CGCGCGGCGGCGGCGCTTCGCCAGTGCAGCGGCAAGCGGCACGCGCAAGAAGTCC 1080
DB 1432 CGCGGGAAGCGCGGCGCCATTCGGCCATTCGCCAGGCAAGCGGCAAGCGGCCAGCAAGAACCTC 1491

QY 1081 AGGCTACGCTGCAGCAAGAACGCCCTCGACGTCGAACTTCAAGGAGCTGGGCTGGGACGAC 1140
DB 1492 AAGGCTCGCTGCAGTCGCAAGGCTTTCATGTCAACTTCAAGGACATGGGCTGGGACGAC 1551

QY 1141 TGGATTATCGCGCCCTCGAGTAGGAGGCTATCATCTGGAGGGGTATGCGGACTTCCCG 1200
DB 1552 TGGATCATCGCACCTCTTGAGTATGAGGCTTCCACTGGCGAAGGACTGTGTGAGTTCCCC 1611

QY 1201 CTGCGCTCGCACCTGGAGCCCCCAACCGCCATCATCCAGACGCTGATGAACTCCATG 1260
DB 1612 TTGCGCTCCCACTTGGAGGCCCAACACCGCAGCTCATTCAGACCCCTAATGAACTCTATG 1671

QY 1261 GACCCGGGCTCCACCGCGCCAGCTGCTGCGTGCCGCCCAAAATTGACTCCCATCAGCAAT 1320
DB 1672 GACCCCTGAATCCACACACCCACTTGTGTGCTTACAGGCTGAGTCTATTAGCATC 1731

QY 1321 CTATACATGACCGGGCAATATATGCTGTCTACAAGCAGTACAGGACATGCTGGTGGAG 1380
DB 1732 CTCTTCATGACTCTGCGCAACACGTTGTTATAAACAGTACGAGGACATGCTGCTGGAA 1791

QY 1381 TCGTGGCTGACAGTAGGCGTTCCTTTCCGCGCGCTT 1419
DB 1792 TCTTGTGGCTGCAGTAGCAGCACCGCGCCACCTGTCTT 1830

RESULT 15
US-09-833-381-582
; Sequence 582, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIORITY FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 582
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-582

Query Match      16.6%; Score 237; DB 10; Length 475;
Best Local Similarity 74.1%; Pred. No. 7.5e-40;
Matches 300; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 996 CAGGCGCTTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1055
DB 8 CAGACGAGGAGGACGGCGTGTGGCGGAGCGGAGCATCGAGGCGCGCGCGCGCGCGCGCGCGCG 67

QY 1056 TGGCAAGCGCGCACGGCAAGAAGTCCAGGCTACGCTGCAGCAAGAAGCCCTCGCACGTGAA 1115
```

Search completed: November 25, 2002, 02:53:58
Job time : 102 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 01:08:25 ; Search time 82 Seconds
(without alignments)
739.378 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNVVYKQYEDMVVESCGR 455

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2447	100.0	455	22	AAW50216 Human growth/diffe
2	2447	100.0	455	23	AAU79173 Human growth/diffe
3	2437	99.6	455	23	AAE17604 Human extracellular
4	1798	73.5	436	17	AAW95636 Cartilage-derived
5	1747	71.4	321	16	AAW78730 Human mature VL-1
6	1747	71.4	321	18	AAW26591 Human bone morphog
7	1747	71.4	321	22	AAE10982 Human full length
8	1207.5	49.3	263	16	AAW78739 Murine mv2 protein
9	1207.5	49.3	263	18	AAW26595 Murine BMP-13 homo
10	1202.5	49.1	263	22	AAE10985 Murine partial mv2

11	927.5	37.9	501	16	AAW69600 New TGF-beta famil
12	927.5	37.9	501	18	AAW36100 Human MP52. Homo
13	927.5	37.9	501	18	AAW19210 Human TGF-beta pro
14	927.5	37.9	501	18	AAW11900 Human high mol. wt
15	927.5	37.9	501	18	AAW01799 Human MP52 protein
16	927.5	37.9	501	18	AAW12770 Human bone morphog
17	927.5	37.9	501	19	AAW44868 TGF-beta superfam1
18	927.5	37.9	501	19	AAW33008 Human MP52. Homo
19	919.5	37.6	495	15	AAW60022 Growth differentia
20	919.5	37.6	495	22	AAW84550 Amino acid sequenc
21	918.5	37.5	501	17	AAW95635 Cartilage-derived
22	918.5	37.5	501	21	AAW92034 Human growth diffe
23	917.5	37.5	401	14	AAW40800 TGF-beta-like clon
24	916.5	37.5	501	22	AAW70529 Human TGF-beta MP5
25	765	31.3	388	16	AAW78734 Human bone morphog
26	765	31.3	388	18	AAW26592 Human bone morphog
27	765	31.3	388	22	AAE10986 Human bone morphog
28	731	29.9	134	16	AAW66867 GDF-6. Mus sp. A
29	731	29.9	134	21	AAW12986 Murine growth diff
30	691.5	28.3	294	16	AAW78729 Human bone morphog
31	691.5	28.3	294	18	AAW26589 Human bone morphog
32	691.5	28.3	294	22	AAE10972 Human bone morphog
33	666	27.2	120	23	AAW51933 Murine TGFbeta pro
34	627	25.6	411	16	AAW78740 Fusion of BMP-2 pr
35	627	25.6	411	18	AAW26597 BMP-2 propeptide/B
36	627	25.6	411	22	AAE10983 Human BMP2 propept
37	618	25.3	240	16	AAW78738 Murine mv1 protein
38	618	25.3	240	18	AAW26594 Murine BMP-12 homo
39	618	25.3	240	22	AAE10984 Murine partial mv1
40	601	24.6	184	22	ABG29375 Novel human diagno
41	575	23.5	102	21	AAW09554 Human CDMP-2/GDF-6
42	575	23.5	102	21	AAW02820 Human CDMP-2/GDF-6
43	575	23.5	102	21	AAW92579 CDMP-2/GDF-6 finge
44	572	23.4	102	21	AAW09555 Murine GDF-6, SEQ
45	572	23.4	102	21	AAW02821 Mouse GDF-6 amino

ALIGNMENTS

RESULT 1				
AAW50216				
ID	AAW50216 standard; Protein; 455 AA.			
AC	AAW50216;			
XX				
DT	07-JAN-2002 (first entry)			
XX				
DE	Human growth/differentiation factor-6-like protein AMF10.			
XX				
KW	AMF10; human; growth/differentiation factor-6; cancer;			
XX				
FT	cell proliferation; astrocytoma; glioma; therapy; diagnosis.			
OS	Homo sapiens.			
XX				
FH	Key			
FT	Peptide			
FT	Protein			
FT	FT			
FT	FT			
PN	WO200174897-A2.			
XX				
PD	11-OCT-2001.			
XX				
PF	03-APR-2001; 2001WO-US10892.			
XX				
PR	03-APR-2000; 2000US-194314P.			
XX				
PA	(CURA-) CURAGEN CORP.			
XX				
PI	Vernet CAM, Burgess CE, Fernandes E, Taupier RJ, Quinn KE;			

PI	Sytek KA, Rastelli L, Herrmann JL;
XX	
DR	WPI; 2001-626395/72.
DR	N-PSDB; AAI70203.
XX	
PT	New AMF1-10 polypeptides and encoding polynucleotides, useful for
PT	treating or preventing disorders related to modulation of cell
PT	movement, cell signal processing, cell adhesion or migration pathways
PT	e.g., cancer
XX	
PS	Claim 1; Page 44-45; 134pp; English.
PS	
XX	
CC	The present sequence is that of the novel, secreted human
CC	growth/differentiation factor-6 (GDF6)-like protein, AMF10.
CC	AMF10 is expressed in astrocytoma and glioma-derived tissue. DNA
CC	encoding the AMF8 may be useful in gene therapy, and the protein
CC	may also be used as a therapeutic, especially in treatment of
CC	cancer and other cell proliferative disorders. Generally, the
CC	AMF1-10 (AMFX) nucleic acids and proteins of the invention are useful
CC	for treating or preventing AMFX-associated disorders, e.g. a disorder
CC	related to cell signal processing and metabolic pathway modulation,
CC	cell adhesion or migration pathway modulation, chemoresistance,
CC	radiotherapy resistance, survival in trophic factor limited
CC	secondary tissue site microenvironments, connective tissue
CC	disorders, tissue remodeling, oncogenesis, cancer of the breast,
CC	ovary, cervix, prostate, endometrium, stomach, colon, lung,
CC	bladder, kidney, brain, and soft-tissue, cellular transformation,
CC	developmental tissue remodeling, inflammation, blood clot
CC	formation and resorption, haematopoiesis, angiogenesis, multidrug
CC	resistance related to organic anion transporters, malignant disease
CC	progression, autocrine and paracrine regulation of cell growth, and
CC	cellular responses to external stimuli, and other diseases,
CC	disorders, etc. (all claimed). AMFX proteins are also used for
CC	screening drugs or compounds that modulate AMFX protein activity or
CC	expression as well as to treat disorders characterized by
CC	insufficient or excessive production of AMFX protein.

Db	421	VPTKLTPTISILYIDAGNNVYKQVEDMWVESCGR	455
RESULT 2			
AAU79173			
ID	AAU79173	standard; Protein; 455	AA.
XX	AC		
XX	AAU79173;		
XX			
DT	02-JUL-2002	(first entry)	
XX			
DE	Human growth/differentiation factor 6-like protein NOV9.		
XX			
KW	Human; growth/differentiation factor 6-like protein; NOVX; NOV9;		
KW	vaccine; cancer; tumour; bone disorder; avascular necrosis; allergy;		
KW	haematopoietic disorder; immune disorder; endometriosis; renal disease;		
KW	infection; inflammatory disease; lung disease; scleroderma; ataxia;		
KW	bowel disease; appendicitis; blood disorder; cardiovascular disorder;		
KW	graft versus host disease; GVHD; lymphoedema; brain disorder;		
KW	ocular disorder; hepatitis C virus infection; cardiac disorder;		
KW	autosomal dominant deafness; DFNA-2.		
XX			
OS	Homo sapiens.		
XX			
PN	W0200214368-A2.		
XX			
PD	21-FEB-2002.		
XX			
FF	16-AUG-2001; 2001WO-US25624.		
XX			
PR	16-AUG-2000; 2000US-225692P.		
PR	16-AUG-2000; 2000US-225693P.		
PR	16-AUG-2000; 2000US-225837P.		
PR	18-AUG-2000; 2000US-226236P.		
PR	18-AUG-2000; 2000US-226236P.		
PR	22-AUG-2000; 2000US-227085P.		
PR	23-AUG-2000; 2000US-227395P.		
PR	24-AUG-2000; 2000US-227492P.		
PR	24-AUG-2000; 2000US-227600P.		
PR	14-MAR-2001; 2001US-275952P.		
XX			
FA	(CURA-) CURAGEN CORP.		
XX			
PI	Zerhusen BD, Padigaru M, Spytek KA, Spaderna SK, Ganqolli EA;		
PI	Rastelli L, Burgess CB, Majumder K, Shimkets R, Mishra V;		
PI	Vernet CAM, Szekeres ES, Grosse WM, Alsobrook JP, Liu X;		
PI	Gerlach WL, Ellerman K, Smithson G, Peyman J, Stone D;		
PI	MacDougall J;		
XX			
DR	WPI; 2002-329571/36.		
DR	N-PSDB; ABK48393.		
XX			
PT	Novel cytoplasmic, nuclear membrane bound and secreted NOVX		
PT	polypeptides, useful for treating cancers and tumours, bone disorders,		
PT	Paget's disease, haematopoietic disorders, spinal diseases and immune		
PT	disorders		
XX			
PS	Claim 1; Page 86; 234pp; English.		
XX			
CC	The present invention relates to new isolated NOVX polypeptides named		
CC	NOV1-NOV9. The invention can be used for identifying an agent (a cellular		
CC	receptor or downstream effector) that binds to the polypeptide. The		
CC	molecules of the invention are useful for treating or preventing		
CC	NOVX-associated disorders in humans. The antibody of the invention is		
CC	useful for determining the presence or amount of NOVX in a sample, and		
CC	for treating a pathological state in a mammal. The method of the		
CC	invention is useful for determining the presence of an amount of NOVX in		
CC	a sample which is used as a marker for cancerous cell or tissue type.		
CC	The molecules of the invention are useful in the manufacture of a		
CC	medicament for treating or preventing cancer, tumour, bone disorders,		
CC	avascular necrosis, allergy, haematopoietic disorders, immune disorders,		
CC	endometriosis, renal diseases, infections, inflammatory diseases, lung		
CC	diseases, scleroderma, ataxia, bowel diseases, appendicitis, blood		

06-JUN-2000; 2000US-210233P.
23-JUN-2000; 2000US-213455P.
14-NOV-2000; 2000US-249019P.
(INCY-) INCYTE GENOMICS INC.

Lal P, Yue H, He A, Nguyen DB, Wallia N, Gandhi AR, Azimzai Y, Bandman O, Tang YT, Lu Y, Baughn MR, Duggan BM, Lee S, Hafalia J, Policy JL; WPT; 2002-154573/20.

NY 10027, NY 10027, NY 10027.

new extracellular messenger polypeptides and polynucleotides encoding them, useful for diagnosing, treating or preventing e.g. neurological,

and polynucleotides encoding them. XMEs is useful in the diagnosis,

stroke, or Alzheimer's disease), autoimmune/inflammatory disorders (e.g., acquired immune deficiency syndrome, AIDS), Addison's disease, or

allergies), developmental disorders (e.g. renal tubular acidosis, anaemia or Cushing's syndrome), endocrine disorders (e.g. hypophysectomy

aneurysm or vascular malformation), and cell proliferative disorders (e.g. cancer), and in the assessment of the effects of exogenous

compounds on the expression of nucleic acid and amino acid sequences of extracellular messengers. XMEs may also be used in the treatment of

viral, bacterial, fungal, parasitic, protozoal and helminthic infections, trauma, disorders associated with hypopituitarism,

hypochloridism, hyperthyroidism or gonadal steroid hormones, and pancreatic disorders such as type I or type II diabetes mellitus. The

may be used for germ-line gene therapy. The present sequence is human XMES-6 protein.

Sequence 455 AA;

very Match 99.68; Score 2437; DB 23; Length 455;

Matches	454;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps
---------	------	--------------	----	------------	----	--------	----	------

1 MDTPRVLLSAVFLISFLWDLPGFQQASISSSCSSAELGSTKGMRSRKEGKMQRAPRDSA 60

1 MDTPRVLLSAVFLISFLWDLPGFQQASISSSSSSAELGSTKGMRSRKEGKMQRAPRDSA 60

61 GREGQEPQPRPQDEPRAQQPRAQEPGRCGRPVVPHEYMLSIYRTYSIAEKLGINASFFQS 120

61 GREGQEPQPRPQDEPRAQQPRAQEPGPPGPRVVPHEYMLSLYRTYSIAEKLGINASFFQS 120

IZI SKSANITTSFVDRGLDLSHIPLRKRQKYLFFDVSMLSDKREELVGAELRLEFRQAPSAPWGPP I80

121 0K3AN1115FVDRGDDDSHIFLRKQK11LF DV3MLESDAKEELVGAEDKLE KQAF3AFWGF 100

[illegible]

1. The first step in the process is to identify the problem or issue that needs to be addressed. This involves gathering information and understanding the context of the problem.

1000

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

361 VNEKEI CLEDDWITARI EYEAHVCEVCDEEDIPCHI EEDNHATOTI MNCMDPCCSTPDSOC 420

421 VPMKLTPTSTLYIDACNNMNVKQVFDMVVECCCP 155


```

RESULT 7
AAE10982
ID AAE10982 standard; Protein; 321 AA.
XX
AC AAE10982;
XX
XX 18-DEC-2001 (first entry)
XX
XX Human full length VL-1 or BMP-13 protein.
XX
KW Human; bone morphogenic protein; BMP-13; vulnerary; antiinflammatory;
KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;
KW tissue formation; wound healing; tissue repair; ligament defect; VL-1;
KW carpal tunnel syndrome; tendonitis.
XX
XX Homo sapiens.
XX
FH Key
FT Peptide
FT /label= Signal_peptide
FT Protein
FT /label= Mature_BMP_12_protein
XX
XX US6284872-B1.
XX
XX 04-SEP-2001.
XX
XX 28-FEB-1997; 97US-0808324.
XX
XX 22-DEC-1994; 94US-0362670.
XX 07-DEC-1993; 93US-0164103.
XX 25-MAR-1994; 94US-0217780.
XX 02-NOV-1994; 94US-0333576.
XX
XX (GEMY ) GENETICS INST INC.
XX (HARD ) HARVARD COLLEGE.
XX
XX Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;
XX Melton DA;
XX
XX WPI; 2001-588978/66.
XX N-PSDB; AAD18333.
XX
XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel
XX syndrome and other tendon and ligament defects, comprises DNA encoding
XX propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
XX BMP-13 or MP52 -
XX
XX Example 1; Column 57-60; 42pp; English.
XX
XX The invention relates to a chimeric DNA comprising a DNA sequence
XX encoding a propeptide from a member of the transforming growth factor
XX (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
XX encoding an amino acid sequence encoding a mature polypeptide consisting
XX of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
XX sequences are useful for producing proteins which induce tendon/ligament
XX like tissue formation, and for isolating and cloning further DNA
XX sequences encoding BMP-12 related proteins with similar activity. The
XX proteins are useful for the induction of tendon/ligament-like tissue
XX formation, wound healing, ligament and other tissue repair, augmenting
XX the activity of bone morphogenetic proteins, and for treating tendonitis,
XX carpal tunnel syndrome and other tendon and ligament defects. The
XX present sequence is human VL-1 protein also designated as BMP-13.
XX
XX Sequence 321 AA:
XX
XX Query Match 71.4%; Score 1747; DB 22; Length 321;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-140;
XX Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 137 DLSHTPLRQKYLFDVSMLSKDEELVGAELRFRQAPSPAGPLHVLQFPCLSPLL 196
XX

```

PR 22-DEC-1994; 94US-036267

DR N-P

WPI; 1997-424270/39.

BMP-13; bone morphogenetic protein; mouse; tendon; ligament;

```
PR 07-DEC-1993; 93US-0164103.
PR 25-NAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX
PA (GENY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
PI Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;
PI Melton DA;
XX
XX WPI: 2001-588978/66.
DR N-PSDB; AAD18336.
XX
XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel
PT syndrome and other tendon and ligament defects, comprises DNA encoding
PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
PT BMP-13 or MP52 -
XX
XX Example 1; Column 71-74; 42pp; English.
PS
XX The invention relates to a chimeric DNA comprising a DNA sequence
CC encoding a propeptide from a member of the transforming growth factor
CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
CC encoding an amino acid sequence encoding a mature polypeptide consisting
CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
CC sequences are useful for producing proteins which induce tendon/ligament
CC like tissue formation, and for isolating and cloning further DNA
CC sequences encoding BMP-12 related proteins with similar activity. The
CC proteins are useful for the induction of tendon/ligament-like tissue
CC formation, wound healing, ligament and other tissue repair, augmenting
CC the activity of bone morphogenetic proteins, and for treating tendonitis,
CC carpal tunnel syndrome and other tendon and ligament defects. The
CC present sequence is murine partial mv2 protein which is homologous to
CC human BMP-12 or VL-1 sequences of the invention.
XX
XX Sequence 263 AA;
SQ
Query Match 49.1%; Score 1202.5; DB 22; Length 263;
Best Local Similarity 83.3%; Pred. No. 4.3e-94;
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;
QY 180 PAGPLHVQLFPCLLDLDARTLDQAGAPPAGWEVDVWQGLRHQPWQLCLELRAANGE 239
DB 7 PAG-----PTLRGSSQTQPR---PAG-KSFVDWQGLRPQFWQLCLELRAANGE 51
QY 240 LDAGEAARAGQPQPPDLRLSLGFRVRPPQERALLVFTRSQRNLFAMREQLGS 299
DB 52 LDXTGARAGQPQPPDLRLSLGFRVRPPQERALLVFTRSQRNLFTEMHEQLGS 111
QY 300 ABAAGPGAGAGSWPPPGAPDARPLPSGRRRRRTAFASRHKRHKKSRLCRSKPL 359
DB 112 AEA---AGAEGSCAPSGSDTGSWLPSPGRRRRRTAFASRHKRHKKSRLCRSKPL 167
QY 360 HYNFKELGWDWIIAPLEYAHCVCVDFPLRSHLEPTNHAIITLNMNSMDPGSTPPSC 419
DB 168 HYNFKELGWDWIIAPLEYAHCVCVDFPLRSHLEPTNHAIITLNMNSMDPGSTPPSC 227
QY 420 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 455
DB 228 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 263
RESULT 11
ID AAR69600
XX AAR69600 standard; Protein; 501 AA.
XX AAR69600;
AC AAR69600;
XX
DT 10-OCT-1995 (first entry)
DE New TGF-beta family member - MP-52 protein sequence.
XX Transforming growth factor-beta family; mitogenic; differentiation;
KW treatment; prevention; disease; bone; cartilage; connective tissue;
KW skin; mucosa; epithelium; dental tissue; wound healing; osteoporosis;
KW tissue regeneration; arthritis; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 382
FT /label= mature protein
XX
XX WO9504819-A.
XX
XX 16-FEB-1995.
XX
XX 09-AUG-1994; 94WO-EP02630.
XX
XX 10-AUG-1993; 93DE-4326829.
XX 25-MAY-1994; 94DE-4418222.
XX 09-JUN-1994; 94DE-4420157.
XX
XX (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.
XX
XX Hotten G, Neidhardt H, Paulista M, Hoetten G;
XX WPI: 1995-090897/12.
XX N-PSDB; AAQ83695.
XX
XX New DNA encoding a new member of the TGF beta family - and
PT related vectors, host cells etc., has mitogenic and
PT differentiation inducing activity, e.g. for treating or
PT preventing diseases of bone and cartilage etc.
XX
XX Claim 6; Page 36; 51pp; German.
XX
XX The amino acid sequence of a novel member of the transforming growth
CC factor-beta (TGF-b) family named MP-52. The gene encodes a protein
CC of 501 amino acids (AA). The protein, or at least the mature protein,
CC has mitogenic and/or differentiation inducing properties useful in
CC the treatment or prevention of diseases of bone, cartilage, connective
CC tissue, skin, mucosa, epithelium or dental tissue. The protein can also
CC be used for wound healing and tissue regeneration e.g. in osteoporosis
CC and arthritis.
XX
XX Sequence 501 AA;
SQ
Query Match 37.9%; Score 927.5; DB 16; Length 501;
Best Local Similarity 44.7%; Pred. No. 2.3e-70;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;
QY 38 GSTKGMRSRKEGKMQRAPRDSADRGEGQEPQPRQDEPRAQQ----- 79
DB 81 GQTGGLTQPKDEPKLP-----PRPG-GPEPKGHPQTRQATARTVTPKQLPGGKAP 134
QY 80 -----PRAQEPGGRGPRVPHYEMLSITYRSTYSAEIKGINASFPQ 119
DB 135 PKAGSVSPSSFLKKAREPGPPREPKEPFPPTITPHEYMLSLYRTLSDADKGGNSVKL 194
QY 120 SSKSANTITSFVDRGLDLSHTPLRROKYLFDVMSLSKEELVGAELRFRQAPS----- 174
DB 195 EAGLANTITSFIDKGDORGPV-VRKQRYVFDISAL-EKDGLLGAELRLILRKRPSTAKP 252
QY 175 -APWGPAGPLHVQLFPCLS-----PLLDARTLDQAGAPPAGWEVDVWQGLRH-QPWKQ 228
DB 253 AAPGGGAAQL--KLSSCPSGRQPSALLDVRSV--PGLDGGSGWEVDIWKLFNFKNSAQ 308
QY 229 LCLELRAAWGELDAGEAEARAGQPQPPPPDLRSIGFGRVRPPQERALLVFTTSQRKN 288
DB 309 LCLELE-AW---ERGRA-----VDLURLGLDFDRAARQVHEKALFLVFGRTKKRD 352
QY 289 L-FAEMREQLGSAAEAGPGAGAGSWPPPGAPDARPLPSGRRRRRTAFASRHKRHK 347
DB 353 LFFNEIKARSGQDDKT-----VVEYLFQ-RRRRRAPLATROGKRPS 393
```

```
PR 07-DEC-1993; 93US-0164103.
PR 25-NAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX
PA (GENY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
PI Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;
PI Melton DA;
XX
XX WPI: 2001-588978/66.
DR N-PSDB; AAD18336.
XX
XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel
PT syndrome and other tendon and ligament defects, comprises DNA encoding
PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
PT BMP-13 or MP52 -
XX
XX Example 1; Column 71-74; 42pp; English.
PS
XX The invention relates to a chimeric DNA comprising a DNA sequence
CC encoding a propeptide from a member of the transforming growth factor
CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
CC encoding an amino acid sequence encoding a mature polypeptide consisting
CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
CC sequences are useful for producing proteins which induce tendon/ligament
CC like tissue formation, and for isolating and cloning further DNA
CC sequences encoding BMP-12 related proteins with similar activity. The
CC proteins are useful for the induction of tendon/ligament-like tissue
CC formation, wound healing, ligament and other tissue repair, augmenting
CC the activity of bone morphogenetic proteins, and for treating tendonitis,
CC carpal tunnel syndrome and other tendon and ligament defects. The
CC present sequence is murine partial mv2 protein which is homologous to
CC human BMP-12 or VL-1 sequences of the invention.
XX
XX Sequence 263 AA;
SQ
Query Match 49.1%; Score 1202.5; DB 22; Length 263;
Best Local Similarity 83.3%; Pred. No. 4.3e-94;
Matches 230; Conservative 6; Mismatches 21; Indels 19; Gaps 4;
QY 180 PAGPLHVQLFPCLLDLDARTLDQAGAPPAGWEVDVWQGLRHQPWQLCLELRAANGE 239
DB 7 PAG-----PTLRGSSQTQPR---PAG-KSFVDWQGLRPQFWQLCLELRAANGE 51
QY 240 LDAGEAARAGQPQPPDLRLSLGFRVRPPQERALLVFTRSQRNLFAMREQLGS 299
DB 52 LDXTGARAGQPQPPDLRLSLGFRVRPPQERALLVFTRSQRNLFTEMHEQLGS 111
QY 300 ABAAGPGAGAGSWPPPGAPDARPLPSGRRRRRTAFASRHKRHKKSRLCRSKPL 359
DB 112 AEA---AGAEGSCAPSGSDTGSWLPSPGRRRRRTAFASRHKRHKKSRLCRSKPL 167
QY 360 HYNFKELGWDWIIAPLEYAHCVCVDFPLRSHLEPTNHAIITLNMNSMDPGSTPPSC 419
DB 168 HYNFKELGWDWIIAPLEYAHCVCVDFPLRSHLEPTNHAIITLNMNSMDPGSTPPSC 227
QY 420 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 455
DB 228 CVPTKLTPTISILYIDAGNNVYKQYEDMVVESCGR 263
RESULT 11
ID AAR69600
XX AAR69600 standard; Protein; 501 AA.
XX AAR69600;
AC AAR69600;
XX
DT 10-OCT-1995 (first entry)
DE New TGF-beta family member - MP-52 protein sequence.
XX Transforming growth factor-beta family; mitogenic; differentiation;
```

```
QY 348 KKSRLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSHLEPTNHAIQTLM 407
      | : |||:| |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 394 KNLKARCSKALHVNFKDMGWDWIIAPLEYEAFHCEGLCEPFLRSHLEPTNHAIQTLM 453
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 408 NSMDPGSTPPSCVPTKLPISILYIDAGNNVYKQYEDMVVESCGR 455
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESCGR 501
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
AAW36100
ID AAW36100 standard; Protein; 501 AA.
XX
AC AAW36100;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human MP52.
XX
KW Bone morphogenetic protein; BMP; processing enzyme; MP52;
KW BMP-2; BMP-4; BMP-6; BMP-7; bone formation; bone regeneration.
XX
OS Homo sapiens.
XX
PN WO9741250-A1.
XX
PD 06-NOV-1997.
XX
PF 28-APR-1997; 97WO-JP01474.
XX
PR 30-APR-1996; 96JP-0130618.
XX
PA (FARH ) HOECHST YAKUHIN KOGYO KK.
PA (FARH ) HOECHST PHARM & CHEM KK.
XX
PI Kimura M, Makishima F, Takahashi M;
DR WPI; 1997-549748/50.
DR N-PSDB; AAT98191.
XX
PT Production of mature bone morphogenetic protein - by treatment of
PT precursor protein with a processing enzyme such as furin either
PT directly or by expressing them both in the same host
XX
XX Example 1; Pages 21-25; 34pp; Japanese.
PS
CC The present sequence is MP52, which is a bone morphogenetic
CC protein (BMP).
CC Mature BMP can be produced by directly adding a BMP processing
CC enzyme to a solution containing BMP precursor protein, or by
CC transforming an animal cell with expression vectors containing DNA
CC encoding the enzyme and precursor protein, culturing the
CC transformant and isolating the mature BMP from the culture. The
CC method can be used to produce MP52, BMP-2, BMP-4, BMP-6 and BMP-7,
CC which can be used to treat bone formation or regeneration
CC abnormalities.
XX
SQ Sequence 501 AA;

Query Match 37.9%; Score 927.5; DB 18; Length 501;
Best Local Similarity 44.7%; Pred. No. 2.3e-70;
Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;

QY 38 GSTKGRSRKCKMORAPRDSAGREGQPPRPODEPRAQ----- 79
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 81 GGTGLTPQKDEPKKLP-----PRPG-GPEPKGHPGPTQATARTVTKQLPGGKAP 134
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
QY 80 -----PRAQPPGPRVWPHEYMLSIYRTVSYIAEKLGINASFQ 119
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 135 PKAGSVPSFLKKAREPCPPREPFPPTTHEYMLSIRYLSADRKGSSVKL 194
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
QY 120 SKSANTITSFVDRLGLDLSHTPLRRQRYLFDVMSLSDKEELVGAEILRFQAPS----- 174
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
```

```
Db 195 EAGLANTIITFDKGODDRGPV-VRKQRYVFDISAL-EKDGLLGAEILRLRKKPSDTAKP 252
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
QY 175 -APWGPAGPLHVLQFPCLs----PLLLDARTLDPOGAPPAGWEVFDVWQGLRH-QPWKQ 228
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 253 APGGGAAQL--KLSSCFSGQPASLLDVRV--PGLDGSWEVFDIWKLFNFNSAQ 308
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
QY 229 LCLELRAAWGELDAGEAEARAGPOQPPDLRLSLGFRVRPPQBRALLVVFTRSORKN 288
      ||||| |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 309 LCLELE-AW---ERGRA-----VDLRLGLGFDRAARQVHEKALFLVFGRTKKRD 352
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
QY 289 L-FAEMREQLGSAEAGPCAGAGSGWPPSPGAPDARWLPSPGRRRRRTAFASRHGRHG 347
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 353 LFFNEIKARSGQDDKT-----VVEYLFsq-RRKRRAPLATRQGRKPS 393
      |||:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
QY 348 KKSRLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSHLEPTNHAIQTLM 407
      | : |||:| |||||:|||||:|||||:|||||:|||||:|||||
Db 394 KNLKARCSKALHVNFKDMGWDWIIAPLEYEAFHCEGLCEPFLRSHLEPTNHAIQTLM 453
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 408 NSMDPGSTPPSCVPTKLPISILYIDAGNNVYKQYEDMVVESCGR 455
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 454 NSMDPESTPTCCVPTRLSPISILFIDSANNVYKQYEDMVVESCGR 501
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
AAW19210
ID AAW19210 standard; Protein; 501 AA.
XX
AC AAW19210;
XX
DT 04-MAR-1998 (first entry)
XX
DE Human TGF-beta protein MP52.
XX
KW Human transforming growth factor-beta; TGF-beta; MP52; superfamily;
KW cartilage; bone inducing activity; inhibit; bone resorption.
XX
OS Homo sapiens.
XX
PN DE19548476-A1.
XX
PD 26-JUN-1997.
XX
PF 22-DEC-1995; 95DE-1048476.
XX
PR 22-DEC-1995; 95DE-1048476.
XX
PA (BIOP-) BIOPHARM GES BIOTECHNOLOGISCHEN ENTWICKL.
XX
PI Bechtold R, Hotten G, Paulista M, Pohl J, Hoetten G;
XX
DR WPI; 1997-333931/31.
DR N-PSDB; AAT69695.
XX
PT Compound containing protein from TGF-beta superfamily - has bone
PT and/or cartilage inducing activity, useful in treatment of, e.g.
PT osteoporosis, bone damage, Paget's disease and osteoarthritis
XX
PS Claim 3; Page 9; 10pp; German.
XX
CC This sequence is the human transforming growth factor (TGF)-beta protein
CC designated MP52. MP52 can be used in a compound of formula (I):
CC A-X(1-20)-B(1-20) (I); A = protein, or fragment, of the TGF-beta
CC superfamily with cartilage and/or bone inducing activity (e.g. MP52);
CC B = 1 or more substituent groups with an affinity to the extracellular
CC matrix, cellular components of bone and/or cartilage and/or to a
CC biocompatible carrier matrix; X = 1 or more covalent bonds and/or spacer
CC groups. The compound may be used to inhibit bone resorption, prevent or
CC treat bone or cartilage related disorders, including osteoporosis,
CC Paget's disease, osteodystrophy, osteoarthritis or osteoarthritis and
CC to treat bone or cartilage damage caused by wounding or overloading.
XX
SQ Sequence 501 AA;
```

	Query Match	37.98;	Score 927.5;	DB 18;	Length 501;
	Best Local Similarity	44.7%;	Pred. No. 2.3e-70;		
	Matches 209;	Conservative	64;	Mismatches	98; Indels 97; Gaps 16;

QY	38	GSTGMRSRKEGMORAPRDSADAGEQEOPROPQDEPRAOQ-----	79	
Db	81	QTGGTLPKNDPEKKLPP---PRPG-GPEPKFGHPHQTRQAARTVTPKGQLPGGKAP	134	
QY	80	-----PRAGEPGRGVVPVPHYYMLSYRTYSIAEKLGINASFQQ	119	
Db	135	PKAGSVPSFLLKKAREGPGRPPEKPFRPPITPHEYMLSLYRTLSDADRKGNGSSVKL	194	
QY	120	SSKSANITTSVDRLGLDDLSHTPLRRKYFLPDVSMLSKEELVGAEALRFROAPS----	174	
Db	195	EAGLANITITSFDKGDQRDPV-VRKQRYVFDISAL-EKDGLLGAEILIRKKPSTDAKP	252	
QY	175	-APWPPAGPLHVOLFCLSL----PLLDDARTLDPOGAPPAGWEVFWWGCLRH-QPWKQ	228	
Db	253	AAPCGGRAAQL--KLSSCPSGRQPASLLDVRSV--PGLDGSGEVEFDIWKLFNRFNKSAQ	308	
QY	229	LCLERLAAMELDAGEAEAARGPQQPPPDLRLSIGFRRVRPPQEARALLVVFTRSORKN	288	
Db	309	LCLELE-AW-----VDLRGLGEDRAARQVHEKALFLVGTGTTKRDR	352	
QY	289	L-FAMEREQLSAAEGAGAGSWPPSPGADARPWLPSGRRRRRTAFASRHGRKHG	347	
Db	353	LFNEIKARSQDDKT-----VYEYLFSQ-RRKRAPLATRGKRPS	393	
QY	348	KKSRLRSCKPLHYNFKELGWDWIAPLEYEAHCBCVCDFPLRSHLEPTNHAIQTLM	407	
Db	394	KNLKARCSRALHYNFKDMGWMDWIAPLEYEAHFCEGLCFPLRSHLEPTNHAVIQTLM	453	
QY	408	NSMPDGSTPFSCCVPTKLTPTSISILIYIDAGNNVYKQVEDMVVESGCGR	455	
Db	454	NSMDPESTPTCCVPTRLSPISILIFIDSANNVYKQEDMVVESGCGR	501	
RESULT 14				
AAW11900				
ID	AAW11900 standard; Protein; 501 AA.			
XX	AAW11900;			
DT	28-OCT-1997 (first entry)			
DE	Human high mol. wt. protein MP52, a growth/differentiation factor.			
KW	Growth factor; differentiation; bone induction; osteoporosis; teeth;			
KW	tooth; dental; joint tissue; cartilage; mucous membrane; skin; nails;			
KW	wound healing; regeneration; skeletal disorder; fracture; dimer.			
OS	Homo sapiens.			
XX	WO9704095-A1.			
PN	06-FEB-1997.			
XX	24-JUL-1996; 96WO-JP02065.			
XX	24-JUL-1995; 95JP-0218022.			
XX	{FARM} HOECHST JAPAN LTD.			
PA	{FARM} HOECHST PHARM & CHEM KK.			
PI	Fujino Y, Kawai S, Kimura M, Matsumoto T, Takahashi M;			
XX	WPI; 1997-132636/12.			
DR	N-PSDB; AAT61412.			
XX				
PT	High molecular weight human MP52 growth or differentiation factor -			
PT	promotes bone induction, is useful for treatment and prevention of			
PT	bone disease			

Claim 1; Page 12-16; 25pp; Japanese.

PS XX	AAMW1900 is a high mol. wt. form of a human growth/differentiation factor MP52. MP52 promotes bone induction and is useful for plastic reconstructive surgery, cosmetic facial treatment, bone transplantation and tooth implantation. It is also useful for the treatment and prevention of disorders of bone formation, bone, cartilage, joint tissue, skin, mucous membranes, nails or teeth; for wound treatment and tissue regeneration; and for the treatment of skeletal disorders and fractures..	
SQ	Sequence 501 AA;	
	Query Match 37.9%; Score 927.5; DB 18; Length 501; Best Local Similarity 44.7%; Pred. No. 2.3e-70; Matches 209; Conservative 64; Mismatches 98; Indels 97; Gaps 16;	
QY	38 GSTGMSRRKEGMQRAPROSDAGREGQEPPRQPDEPRAQQ-----79 : : : : : : : : : : : :	
DB	81 QGTGLTQPKKPEKKLP----PRPG-GEPKGPHGPOTRTARTVTFKGPLGGKAP 134 : : : : : : : : : : : :	
QY	80 -----PRAQEGGRGVVPVHEMYLSIVGYTSIAEKLGINASFFQ 119 : : : : : : : : : : : :	
DB	135 PKAGSVSSFLKKAREGPPREPREFRPPTTPHMYLSTLTSLDADRKGNSSVKL 194 : : : : : : : : : : : :	
QY	120 SSKSANTITSFVDRLGDLDLGHTLPLLRKYLFDSMLSCKEEELGAELRLFQAQS-----174 : :: : : : : : : : :	
DB	195 EAGLANITISFDIKGDDRGPV-VRKQRYVEDISAL-EKDGLLGAEILRIURKKPSDTAKP 252 : :: : : : : : : : :	
QY	175 -APWGPPAGLIHVOLFCLSL-----PLLIDARTLDQGAPPAGEVFVVWQLRH-QPWKQ 228 : : : : : : : : : : : :	
DB	253 AAPGGGRAAQL--KLSCSPSGROPASLLDVRSV--PGLDGSGWEVEFIMKLFRFNNSAQ 308 : : : : : : : : : : : :	
QY	229 LCLELAHWGELEDAEAERAGPQPDPDLRLSLGFGRVRVPPOERALLVVTRSORKN 288 :: : : : : : : : :	
DB	309 LCLEE-AW---EKGRA-----VDRLGLEEDRAARQHKAULFLVGERTAKRD 352 : : : : : : : : : : : :	
QY	289 L-FAMEREQIGSAEAGFGAGAEGSWPPPSPGADARPWLSPGRRRRRTAFSRHGKRHG 347 : : : : : : : : : : : :	
DB	353 LFFEINARKSQDDKT-----VYEYLFNQ-RKKRRRAPLATRQGRKPS 393 : : : : : : : : : : : :	
QY	348 KKSRLRCSKPVLHNVELGWDMWIATPLEYAHCBGVCDFFPLRSHLETNHAIIQTLM 407 : :: : : : : : : : :	
DB	394 KNLKARGSRKALHNFNDMGWDMDIATPLEYAFAHCBLCEFPFLRSHLETNHAIIQTLM 453 : :: : : : : : : : :	
QY	408 NSMDDGSTPPCCCVTKLTPISTILYIDAGNNVYKYEDMVVESCGCR 455 :: : : : : : : : :	
DB	454 NSMDDPESTPTCCVPTLSPISILFIIDSANNVYKYEDMWVESCGCR 501 :: : : : : : : : :	

RESULT 15
AAW01799
ID AAW01799 standard; Protein; 501 AA.
XX AAW01799;
XX AC
XX AC
DT 15-OCT-1997 (first entry)
XX Human MP52 protein.
DE Homo sapiens.
KW Human; MP52; transforming growth factor; TGF; beta; medicament;
KW treatment; prevention; nervous system; disease; neuropathology;
KW ageing.
XX OS
XX DE19525416-A1.
PN PN
PD 16-JAN-1997.
XX XX
PF 12-JUL-1995; 95DE-1025416..
XX PR 12-JUL-1995; 95DE-1025416..

Job time : 84 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 00:59:55 ; Search time 2432 Seconds
(without alignments)
9489.543 Million cell updates/sec

Title: US-09-825-751A-19
Perfect score: 1425
Sequence: 1 ctctcgggagacgagcaca.....tttccgcgcgttgccgcg 1425

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_hic:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_hic:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_nam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	534.2	37.5	543	13	BI559439
2	345	24.2	450	10	BE098696
3	303.4	21.3	403	9	AI535156
4	271.4	19.0	563	12	BR044119
5	253.6	17.8	661	14	BQ169372
6	250.2	17.6	343	12	BF542757

7	237.6	16.7	798	14	BQ604531
8	237	16.6	475	9	AI761284
9	236	16.6	505	13	BI401488
10	230.8	16.2	1025	12	BF588622
11	217.4	15.3	552	9	AI641304
12	209.4	14.7	307	12	BF420157
13	191.8	13.5	479	14	BQ563481
14	191.6	13.4	283	17	AQ311900
15	188.4	13.2	273	10	BE114678
16	186.2	13.1	712	10	BE637633
17	171.4	12.0	597	12	BG162724
18	159	11.2	583	13	BI473851
19	152.4	10.7	608	17	CNS01SYJ
20	144.8	10.2	200	9	AU241555
21	144	10.1	822	13	BI903718
22	144	10.1	841	13	BG696662
23	141.6	9.9	714	12	BF689009
24	139.4	9.8	972	17	CNS0442A
25	138.4	9.7	284	13	BI901630
26	138.4	9.7	316	13	BI901881
27	138.4	9.7	793	12	BG187538
28	138.4	9.7	991	12	BE740547
29	138.2	9.7	525	14	BM691373
30	137.2	9.6	577	12	BF045110
31	136.8	9.6	805	12	BG214178
32	132	9.3	1002	17	CNS04IFG
33	131.4	9.2	656	13	BI400982
34	131	9.2	364	14	BQ567172
35	131	9.2	831	14	BQ715851
36	130.6	9.2	3392	11	BC023526
37	130.2	9.1	563	9	AL788844
38	130.2	9.1	672	14	BQ525130
39	129.6	9.1	855	12	BG186221
40	129.6	9.1	883	9	AL581402
41	129.6	9.1	968	9	AL572764
42	129.2	9.1	817	9	AL576612
43	129	9.1	875	12	BE893005
44	128.4	9.0	441	12	BG385196
45	128.4	9.0	984	9	AI478076

ALIGNMENTS

RESULT 1
BI559439
LOCUS 603252937F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5295441 5',
DEFINITION mRNA sequence.
ACCESSION BI559439
VERSION BI559439.1 GI:15446753
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 543)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabps-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)
DNA Sequencing by: Inceye Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM11747 row: 1 column: 10
High quality sequence stop: 532.
Location/Qualifiers

QY 814 AGTCTGGCTTCGGCGGAGGGTGGCGCC 842
 Db 30 AGTCTGGCTTCGGCGGAGGGTGGCGCC 2

RESULT 3

AI535156/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

AI535156 403 bp mRNA linear EST 03-JUL-1999
 UI-R-C3-sq-b-01-0-UI-s1 UI-R-C3 Rattus norvegicus cDNA clone
 UI-R-C3-sq-b-01-0-UI 3', mRNA sequence.

AI535156

AI535156.1 GI:4449291

EST.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 403)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 3565

Email: msoares@blue.weeg.uiowa.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com) This clone is also available

through the I.M.A.G.E. Consortium at LBNL (info@image.lbnl.gov).

IMAGE ID=1768988

Seq primer: M13 Forward

POLYA=NO.

Location/Qualifiers

1. 403

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C3-sq-b-01-0-UI"

/clone_lib="UI-R-C3"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-C3

library is a subtracted library of a series, ultimately

derived from a mixture of individually tagged normalized

libraries from rat placenta, adult lung, brain, liver,

kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day

embryos, after a series of subtractions to reduce the

representation of cDNAs from which ESTs had already been

generated. The following serially subtracted libraries

were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1

, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5

nucleotides present between the Not I site and the

oligo-dT track which allows identification of the library

of origin of a clone within the mixture. The subtracted

library (UI-R-C3) was constructed as follows: PCR amplified

cDNA inserts from UI-R-C2p clones from which 3' ESTs had

been derived was used as a driver in a hybridization with

the UI-R-C2p library in the form of single-stranded

circles. The remaining single-stranded circles (subtracted

library) was purified by hydroxyapatite column

chromatography, converted to double-stranded circles and

electroporated into DH10B bacteria (Life Technologies) to

generate the UI-R-C3 library. This procedure has been

previously described (Bonaldo, Lennon and Soares, Genome

Research 6:791-806, 1996)"

FEATURES

source

BASE COUNT
 ORIGIN

Query Match 21.3%; Score 303.4; DB 9; Length 403;
 Best Local Similarity 84.6%; Pred. No. 1.3e-49;
 Matches 340; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 441 TCCTCTGCGACACCTCTCCGAGAGACAGAAAGTATTTTGTGTGATGTGTCATGCTCTCAGA 500
 Db 403 TCCTCTGCGACACCTCTCCGAGAGACAGAAAGTATTTTGTGTGATGTGTCATGCTCTCAGA 344
 QY 501 CAAAGAAGAGCTGTGGCGCGGAGCTGCGGCTCTTTCGCCAGCGCCCTCAGAGCCCTG 560
 Db 343 CAAAGAAGAGCTGTGGCGCGGAGCTGAGGCTTATTCGCCAGCGCCCTCAGAGCCCTG 284
 QY 561 GGGGCCACACCGCGCGGCTCCACGTGCAGCTCTTCCCTTGCCTTTCGCCCTTACTGCT 620
 Db 283 GGGGCCACAAACCGACCGCTGCTTTCAGCTCTTCCCTTGTTCCTTTCCTTACT 224
 QY 621 GGACGGCGGACCTCTGACCGCGGCGCGCGCGCGGCTGGGAAGTCTTTCGACGT 680
 Db 223 GGACTTAGGACCTCTGATCTCAGGGCCCAACAGAAGCTGGCTGGGAAGTCTTTCGACGT 164
 QY 681 GTGGCAGGCTCTGGCCACCGCCCTGGAAGCAGCTGTGCTTGGAGCTGCGGGCCGCGATG 740
 Db 163 GTGGCAGGCTCTGGCCCTCAGCGCTGGAAGCAGCTGTGCTTGGAGCTTTCGGCAGCTG 104
 QY 741 GGGCGAGCTGAGCGCGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 800
 Db 103 GGGTGAAGCTGGACGCCAGGAGTTTCAGGAGCGCGCGCGCGCGCGCGCGCGCGCC 44
 QY 801 CCCGGACCTCGGAGCTGTGGCTTTCGGCGGAGGCTGGCGCC 842
 Db 43 CTTGGACCTCGGAGCTGTGGCTTTCGGCGGAGGCTGGCGCC 2

RESULT 4

BF044119/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF044119 563 bp mRNA linear EST 10-OCT-2000
 BP250023A10F2 Soares normalized bovine placenta Bos taurus cDNA
 clone BP250023A10F2 5', mRNA sequence.

BF044119

BF044119.1 GI:10761174

EST.

COW.

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 563)

Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson

J.H.

Bovine ESTs

Unpublished (2000)

Contact: Lewin, H. A.

W. M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Funding for cattle EST sequencing was provided by the USDA National

Research Initiative, Animal Genome Resource Grant AC 99-3205-8534

to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED

from Washington University Genome Center. Vector Trimmi g:

Cross match from Washington University Genome Center PHRAP suite.

This sequence is vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAATACGACTCTACTATAGG

BACKWARD: ATTAACCTCTACTAAG

Insert Length: 563 Std Error: 0.00

Plate: BP250023A10 row: F column: 2

Seq primer: AGCGATAACAATTTTCACACAGGA
High quality sequence stop: 563.

FEATURES

Source
Location/Qualifiers
1..563
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="BP250023A10F2"
/clone_lib="Soares normalized bovine placenta"
/sex="female"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pTT73Pac; Site_1: EcoRI; Site_2: NotI; The cDNA library was contributed by the Soares laboratory and it was constructed and normalized as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806."

BASE COUNT 110 a 177 c 174 g 102 t

ORIGIN

Query Match 19.0%; Score 271.4; DB 12; Length 563;
Best Local Similarity 96.0%; Pred. No. 2.4e-43;
Matches 289; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 1125 GCTGGGCTGGGAGCTGGATTATCGGCCCTGGAGTAGCAGGCTATCATCTGCGAGGG 1184
Db 563 GCTGGGCTGGGAGCTGGATTATCGGCCCTGGAGTAGCAGGCTATCATCTGCGAGGG 504
QY 1185 TGTATGCGACTTCCCGCTCGCCTCGCACCTGGAGCCACCAACCGCCATCATCCAGAC 1244
Db 503 CGTGTGGAGCTTCCCGCTCGCCTCGCACCTGGAGCCACCAACCGCCATCATCCAGAC 444
QY 1245 GCTGATGAATCCATGACGCCCGGCTCCACCGCCCGCAGTGTGGTGGCCACCAAAAT 1304
Db 443 GCTGATGAATCCATGACGCCCGGCTCCACCGCCCGCAGTGTGGTGGCCACCAAAAT 384
QY 1305 GACTCCCATCAGCATTTCTATACATGACGGGGCAATATGTGGTCTACAAGCAGTAGCA 1364
Db 383 GACTCCCATCAGCATTTCTATACATGACGGGGCAATATGTGGTCTACAAGCAGTAGCA 324
QY 1365 GGACATGGTGTGGAGTCGTCGGGCTGACGAGTACGGGCTTCCCGCGGCTTGGCC 1424
Db 323 GGAGATGGTGTGGAGTCGTCGGGCTGACGAGTACGGGCTTCCCGCGGCTTGGCC 265
QY 1425 G 1425
Db 264 G 264

RESULT 5

BQ169372 661 bp mRNA linear EST 01-MAY-2002
LOCUS STR00320 segmentation stage cDNA library Danio rerio cDNA clone
DEFINITION CB278 5' similar to DYNAMO PROTEIN, mRNA sequence.
ACCESSION BQ169372
VERSION BQ169372.1 GI:20376800
KEYWORDS EST.
SOURCE Danio rerio
ORGANISM zebrafish.

REFERENCE BQ169372
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
Loppin,B., Pflumio,S., Steffan,T., Heyer,V., Furthauer,M., Thisse
C. and Thisse,B.
TITLE Expression of the zebrafish genome during embryogenesis (2002)
JOURNAL Unpublished (2002)
COMMENT Other_ESTs: STR00321
Contact: Thisse B
Institut de Genetique et de Biologie Moleculaire et Cellulaire
CNRS, INSERM, ULP
1, rue Laurent Fries, BP163, CU de Strasbourg, 67404 Illkirch Cedex
France
Tel: 33 3 88 65 33 60
Fax: 33 3 88 65 32 01

Email: thisse@igbmc.u-strasbg.fr
EST from a cDNA of a gene whose expression is spatially restricted during embryogenesis. We have established its expression pattern during embryonic development by whole mount in situ hybridization on zebrafish embryos from the gastrula stage to 2 days of development. The corresponding data are available on the zebrafish community database at http://zfinfo.org/cDNA_library_preparation: B. Rigleman. DNA Sequencing by: IGBMC sequencing facility. Clone distribution: zebrafish international resource center at the University of Oregon (Institute of Neuroscience, 1254 University of Oregon, Eugene, OR 97403-1254)
Seq primer: T3 ATTACCTCACTAAAGGA.

FEATURES

Source
Location/Qualifiers
1..661
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="CB278"
/clone_lib="segmentation stage cDNA library"
/dev_stage="segmentation stage embryos"
/note="Vector: Lambda Zap; Site_1: EcoRI; Site_2: XhoI;
Oligo dT cDNA library constructed from RNA pooled from segmentation stage zebrafish embryos"

BASE COUNT 201 a 169 c 172 g 119 t

ORIGIN

Query Match 17.8%; Score 253.6; DB 14; Length 661;
Best Local Similarity 77.5%; Pred. No. 7.6e-40;
Matches 307; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 1005 GCTGCCCTCGCCGGCGCGCGCGCGCGCTTCGCCAGTCGCCAAGCG 1064
Db 198 GCTGCAGTTTAAAGCGCGCGCAGAGAGACTGCGCTTAATAATCGGCACGGAAAG 257
QY 1065 GCAGCGAAGAAGTCAGGCTAGCTGCAGCAGAGAGACCCCTGCAGTGAACCTTCAAGGA 1124
Db 258 GCATGCCAAAAGTCCAAAATCGAGATGCAGCAAAAAGGCTCTGCAGTCAACTTCAAGA 317
QY 1125 GCTGGCTGGGAGCTGGATTATCGGCCCTCGGAGTAGAGGCTATCAGTCGCGAGG 1184
Db 318 ACTGTGAGGGAGCAGCTGGATCATCGCTCCCTGGATTAGAGAGCCCTATCATCTGCGAGG 377
QY 1185 TGTATGCGACTTCCCGCTCGCTCGCAGCTGGAGCCACCAACCGCCATCATCCAGAC 1244
Db 378 CGTGTGCGACTTCCCGTTGAGGTCGCACCTAGAGCGCGACCAACCGCCATCATCCAGAC 437
QY 1245 GCTGATGAATCCATGAGCCCGGCTCCACCGCCCGCGCGCTGCTGCTGCCCAAAAT 1304
Db 438 GCTCATGAATCCATGAGCCCGGCTCCACCGAGCTGTTGCTGCCCAAAACT 497
QY 1305 GACTCCCATCAGCATTTATACATCGACGCGGGCAATAATGTGCTACAAGCAGTAGCA 1364
Db 498 CAGCCCATCAGTATACTGTACATAGACTCTGGGAACACAGCTGCTGTACAACAGTAGCA 557
QY 1365 GGACATGGTGGAGTCTGTCGGGCTGCGAGGTGCGAGGTAGCG 1400
Db 558 GGACATGGTGGTAGAACAGTGTGCTGCGAGGTAGCG 593

RESULT 6

BQ542757 343 bp mRNA linear EST 11-DEC-2000
LOCUS UI-R-C3-sq-b-01-0-UI.r1 UI-R-C3 Rattus norvegicus cDNA clone
DEFINITION UI-R-C3-sq-b-01-0-UI 5', mRNA sequence.
ACCESSION BQ542757
VERSION BQ542757.1 GI:11633864
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 343)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) This clone is also available through the I.M.A.G.E. Consortium at LLNL (info@image.llnl.gov). IMAGE ID= 1768988
Seq primer: M13 Forward.

FEATURES

Location/Qualifiers

1..343

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C3-sq-b-01-0-UI"

/clone_lib="UI-R-C3"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C3 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, and 8, 12 and 18-day embryos, after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: UI-R-C3, UI-R-C2p, UI-R-C1, UI-R-C0, UI-R-A1, UI-R-E1. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within themixture. The subtracted library (UI-R-C3) was constructed as follows: PCR amplified cDNA inserts from UI-R-C2p clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C2p library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C3 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)"

Research 6:791-806, 1996"

BASE COUNT 63 a 99 c 108 g 72 t 1 others

ORIGIN

Query Match 17.6%; Score 250.2; DB 12; Length 343;

Best Local Similarity 83.9%; Pred. No. 3.3e-39;

Matches 282; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGTCATGCTCTCA 498

|||||

Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGTCATGCTCTCA 67

QY 499 GACAAAGAAGAGCTGTTGGCGCGGAGCTCGGCTCTTTCGCCAGCGGCCCTCAGCGCCC 558

|||||

Db 68 GACAAAGAAGAGCTGTTGGCGCGGAGCTGAGGCTTATCGCCAGGCGGCCCAACGCC 127

QY 559 TGGGGGCCACAGCGCGCGCTCCAGTGCAGCTCTTCCTTCCTTCCTTCCTTCCTTCCT 618

|||||

Db 128 TGGGGGCCACAAACCCGACGCTGGATTGTCAGCTCTTCCCTCTGTTGCGCATTTGTA 187

QY 619 CTGGACGCGGAGCCCTTGGACCGCAGGCGCGCGCGCGCTGGGAAGTCTTCGAC 678

|||||

Db 188 CTGGACTCTAGGACCTTGGATCTCTCAGGGGCCAACAGAGCTGCTGGGAAGTTTTCNAC 247

QY 679 GTGTGGCAGGCCTCGCCACCACCGCTTGAAGCAGCTGTCTTGGAGCTCGGGCGGCA 738

|||||

Db 248 GTGTGGCAGGCCTCGCCACCCTCGAAGCAGCTGTCTTGGAGCTTCGGCGCATC 307

QY 739 TGGGCGCAGCTGACGCCGGGGAGGCGCGCGCGC 774

|||||

Db 308 TGGGGTGAGTGGACGCCAGGATTTCAGAGCGCGC 343

RESULT 7

BQ604531

LOCUS BQ604531 798 bp mRNA linear EST 24-JUN-2002

DEFINITION MI-P-CP1-nzc-d-04-0-UI.s1 MI-P-CP1 Sus scrofa cDNA clone

ACCESSION BQ604531

VERSION BQ604531.1 GI:21551257

KEYWORDS EST.

SOURCE Sus scrofa

ORGANISM pig.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 798)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Tuggle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: ktuggle@iastate.edu
Tissue Procurement: Dr. Chris Tuggle, Iowa State University
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-86, >GC-rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD

POLYA=No.

FEATURES

Location/Qualifiers

1..798

/organism="Sus scrofa"

/strain="crossbred"

/db_xref="taxon:9823"

/clone="MI-P-CP1-nzc-d-04-0-UI"

/clone_lib="MI-P-CP1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-CP1 library is normalized library derived from the MI-P-CP0 library, ultimately derived from uterus tissue. For a detailed description of the library from which this clone was derived, please visit our web site at http://pigest.genome.iastate.edu/. The procedure used to create this library has been previously described (Bonaldo , Lennon and Soares, Genome Research 6: 791-806, 1996)

BASE COUNT 169 a 227 c 265 g 137 t

ORIGIN

Query Match 16.7%; Score 237.6; DB 14; Length 798;

Best Local Similarity 74.3%; Pred. No. 1.1e-36;

Matches 300; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 996 CAGGCCTTGGCTCGCCCGCGCGCGCGCGCGCGCGCTTCGCCAGTCCGCA 1055

|||||

Db 8 CAGGCGGTGG 67


```
RESULT 14
A0311900/c
LOCUS
DEFINITION RPCI11-94N16-TV RPCI-11 Homo sapiens genomic clone RPCI-11-94N16,
DNA sequence.
ACCESSION A0311900
VERSION A0311900.1 GI:4043649
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 283)
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPCI11-94N16.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: 17
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..283
/organism="Homo sapiens"
/db_xref="GDB:7536039"
/db_xref="taxon:9606"
/clone="RPCI-11-94N16"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="vector: phage3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 52 a 82 c 83 g 66 t
ORIGIN
Query Match 13.4%; Score 191.6; DB 17; Length 283;
Best Local Similarity 98.0%; Pred. No. 9.7e-28;
Matches 194; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 239 GGCCTCAGACGAACCCGGGCTACGAGCCCGGGCGCAGAGCCGCGAGGGGTC 298
DB 283 GGCCTCAGACGAACCCGGGATCAGCAGCAGCGGGCGCAGGAGCGAGGGGTC 224
QY 299 CGCGGCTGTGTCGCCACGAGTACATGCTCAATCTACAGGACTTACTCCATCGCTGAGA 358
DB 223 CGAGCGTGGTGGCCACGAGTACATGCTCAATCTACAGGACTTACTCCATCGCTGAGA 164
QY 359 AGCTGGGCATCAATGCCAGCTTTTCCAGTCTTCCAAAGTCGAGTCAATACGATCACCAGT 418
DB 163 AGCTGGGCATCAATGCCAGCTTTTCCAGTCTTCCAAAGTCGAGTCAATACGATCACCAGT 104
QY 419 TTGTAGACAGGGGACTAG 436
DB 103 TTGTAGACAGGGGACTAG 86
RESULT 15
BE114678/c
LOCUS
DEFINITION UI-R-BJ1-awo-a-02-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
UI-R-BJ1-awo-a-02-0-UI 3', mRNA sequence.
ACCESSION BE114678
VERSION BE114678.1 GI:8506783
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS 1 (bases 1 to 273)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone Distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-No.
Location/Qualifiers
1..273
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ1-awo-a-02-0-UI"
/clone_lib="UI-R-BJ1"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-BJ1
library is a subtracted library derived from the following
tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
For a detailed description of the library from which this
clone was derived, please visit our web site at
rattest.eng.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_SEQ=None found"
BASE COUNT 54 a 94 c 84 g 40 t 1 others
ORIGIN
Query Match 13.2%; Score 188.4; DB 10; Length 273;
Best Local Similarity 80.8%; Pred. No. 4.1e-27;
Matches 219; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 572 CGGGCCGCTCCACGTGCAGCTTTCCTTGCCTTTCGCCCTACTGTGACGCGCGA 631
DB 272 CCGACCGCTGCATNTGCAGCTTTCCTTTCCTTTCGCCCTACTGTGACTCTAGA 213
QY 632 CCCTGGACCGGAGGGGCGCGCGCGGCTGGGAAGTCTTCAGCTGTGGCAGGGCC 691
DB 212 CCCTGGATCCTCAGGGGCAACAGAGAGCTGGCTGGGAAGTTTTCGACGCTGTGCGAGTCC 153
QY 692 TCGCCACACGCTTGAAGCAGCTGTCTTGGAGCTCGGGCGCATGGGGCGAGCTGG 751
DB 152 TCGCCCTCAGCGGTGAAGCAGCTGTCTTGGAGCTTTCGGCAGCTCTGGGTGAGCTGG 93
QY 752 AGCGCGGGAGGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 811
DB 92 AGCGCAGGATTCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33
QY 812 GGAGTCTGGGCTTCGGCGCGGAGGTGCGGCC 842
DB 32 GGAGTCTGGGCTTCGGCGCGGAGGTGCGGCC 2
```

Search completed: November 25, 2002, 02:50:44
Job time : 2445 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 24, 2002, 23:26:25 ; Search time 310 Seconds
(without alignments)
10351.928 Million call updates/sec

Title: US-09-825-751A-19

Perfect score: 1425

Sequence: 1 ctctcgggagacgagcga.....tttccgcgccttgccgcg 1425

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1425	100.0	1425	22	AAI70203 Human growth/diffe
2	1425	100.0	1425	24	ABK48393 DNA encoding human
3	1421.8	99.8	2791	24	AAAD28496 Human extracellular
4	987	69.3	1171	18	AAAT90386 Human bone morphog
5	987	69.3	1171	22	AAAD18333 Human full length
6	976	68.5	1171	16	AAQ96208 Human mature VL-1
7	928.2	65.1	1308	17	AAAT31602 Cartilage-derived
8	581.8	40.8	1046	18	AAAT90397 Murine BMP-13 homo
9	581.8	40.8	1046	22	AAAD18336 Murine partial mv2

10	581.2	40.8	1046	16	AAQ96224	Murine mv2 protein
11	412.2	28.9	530	16	AAQ84280	GDF-6 gene. Mus s
12	412.2	28.9	530	21	AAA72621	Murine growth diff
13	259.2	18.2	1345	16	AAQ96215	Human bone morphog
14	259.2	18.2	1345	18	AAAT90390	Human bone morphog
15	259.2	18.2	1345	22	AAAD18337	Human bone morphog
16	253	17.8	2323	24	ABQ88241	Human osteoblast d
17	253	17.8	2703	16	AAQ83695	New TGF-beta famil
18	253	17.8	2703	18	AAAT98191	CDNA for human MP5
19	253	17.8	2703	18	AAAT69695	CDNA encoding huma
20	253	17.8	2703	18	AAAT64112	CDNA encoding huma
21	253	17.8	2703	18	AAAT59405	CDNA encoding huma
22	253	17.8	2703	18	AAAT59729	Human bone morphog
23	253	17.8	2703	19	AAAT88340	Human MP52 CDNA.
24	253	17.8	2703	24	ABQ88242	Human osteoblast d
25	251.6	17.7	1207	14	AAQ47709	TGF-beta-like clon
26	251.6	17.7	1207	16	AAQ96209	Murine protein MP5
27	251.6	17.7	1207	22	AAAD18317	Human MP-52 DNA.
28	251.4	17.6	2341	17	AAAT31601	Cartilage-derived
29	250	17.5	2703	22	AAAF74420	Human TGF-beta MP5
30	249.2	17.5	519	16	AAQ82824	GDF-7 C-terminal r
31	245.8	17.2	357	17	AAAT46150	CDNA encoding huma
32	245.8	17.2	357	18	AAAT70296	Human bone inducin
33	245	17.2	1203	16	AAQ96223	Murine mv1 protein
34	245	17.2	1203	18	AAAT90396	Murine BMP-12 homo
35	245	17.2	1203	22	AAAD18335	Murine partial mv1
36	244.8	17.2	926	16	AAQ96207	Human bone morphog
37	244.8	17.2	926	18	AAAT90385	Human bone morphog
38	244.8	17.2	926	22	AAAD18316	Human bone morphog
39	239	16.8	2328	15	AAQ70010	Growth differentia
40	239	16.8	2329	22	AAH28140	Nucleotide sequenc
41	237.8	16.7	357	21	AAAZ29328	Mutant human MP52
42	234.6	16.5	3598	23	AAAS93562	DNA encoding novel
43	232	16.3	272	16	AAQ96221	Murine bone morpho
44	232	16.3	272	22	AAAD18322	Murine clone, mv2
45	230	16.1	1233	16	AAQ96226	Fusion of BMP-2 pr

ALIGNMENTS

RESULT 1
AAI70203
ID AAI70203 standard; DNA; 1425 BP.

XX AAI70203;

XX 07-JAN-2002 (first entry)

DT Human growth/differentiation factor-6-like protein AMF10 DNA.

DE AMF10; human; growth/differentiation factor-6; cancer;
KW cell proliferation; astrocytoma; glioma; gene therapy; diagnosis;
KW ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 31..1398

FT sig_peptide /*tag= a

FT mat_peptide /*tag= b

FT 97..1395

FT /*tag= c

XX WO200174897-A2.

XX 11-OCT-2001.

XX 03-APR-2001; 2001WO-US10892.

XX 03-APR-2000; 2000US-194314P.

XX 16-AUG-2000; 2000US-225693P.

XX (CURA-) CURAGEN CORP.
XX Vernet CAM, Burgess CE, Fernandes E, Taupier RJ, Quinn KE;
PI Spytek KA, Rastelli L, Herrmann JL;
XX WPI; 2001-626395/72.
DR P-PSDB; AAM50216.
XX New AMF1-10 polypeptides and encoding polynucleotides, useful for
PT treating or preventing disorders related to modulation of cell
PT movement, cell signal processing, cell adhesion or migration pathways
PT e.g., cancer
XX Claim 3; Page 44-45; 134pp; English.
XX The present sequence is that of DNA clone AC01136.A. This
CC includes an open reading frame that encodes the novel, secreted
CC human growth/differentiation factor-6 (GDF6)-like protein, AMF10
CC (see AAM50216). AMF10 is expressed in astrocytoma and glioma-derived
CC tissue. DNA encoding the AMF8 may be useful in gene therapy, and the
CC protein may also be used as a therapeutic, especially in treatment of
CC cancer and other cell proliferative disorders. Generally, the
CC AMF1-10 (AMFX) nucleic acids and proteins of the invention are useful
CC for treating or preventing AMFX-associated disorders, e.g. a disorder
CC related to cell signal processing and metabolic pathway modulation,
CC cell adhesion or migration pathway modulation, chemoresistance,
CC radiotherapy resistance, survival in trophic factor limited
CC secondary tissue site microenvironments, connective tissue
CC disorders, tissue remodeling, oncogenesis, cancer of the breast,
CC ovary, cervix, prostate, endometrium, stomach, colon, lung,
CC bladder, kidney, brain, and soft-tissue, cellular transformation,
CC developmental tissue remodeling, inflammation, blood clot
CC formation and resorption, haematopoiesis, angiogenesis, multistep
CC resistance related to organic anion transporters, malignant disease
CC progression, autocrine and paracrine regulation of cell growth, and
CC cellular responses to external stimuli, and other diseases,
CC disorders, etc. (all claimed). Nucleic acid fragments are also
CC useful as probes or PCR primers, for modulating AMFX expression,
CC in chromosome mapping, tissue typing, diagnostics, forensics and
CC pharmacogenomics.
XX SQ Sequence 1425 BP; 234 A; 488 C; 469 G; 234 T; 0 other;
Query Match 100.0%; Score 1425; DB 22; Length 1425;
Best Local Similarity 100.0%; Pred. No. 6.7e-246;
Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTGGGAGAGCGACGACCTTGGCCGATGGATACCTCCAGGGTCTGCTCTCGGCC 60
DB 1 CTCCTGGGAGAGCGACGACCTTGGCCGATGGATACCTCCAGGGTCTGCTCTCGGCC 60
QY 61 GTCCTTCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120
DB 61 GTCCTTCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120
QY 121 TCCCTGTTCTGCGCCGAGCTGGGTTCCACCAAGGGCATCGGAAGCGCGAAGGCAAG 180
DB 121 TCCCTGTTCTGCGCCGAGCTGGGTTCCACCAAGGGCATCGGAAGCGCGAAGGCAAG 180
QY 181 ATGCAGCGGGCGCGCGGACAGTACGACGGCGCGGGAGGGCCAGGAACACAGCCGCGG 240
DB 181 ATGCAGCGGGCGCGCGGACAGTACGACGGCGCGGGAGGGCCAGGAACACAGCCGCGG 240
QY 241 CCTCAGGAGAACCCGGGCTCAGCAGCCCGGGCGGAGGGCCAGGAGCCGAGGGTCCG 300
DB 241 CCTCAGGAGAACCCGGGCTCAGCAGCCCGGGCGGAGGGCCAGGAGCCGAGGGTCCG 300
QY 301 CGCGTGGTCCCGCACAGTACATGCTGTCAATCTACAGGACTTACTTCCATCCTGAGAG 360
DB 301 CGCGTGGTCCCGCACAGTACATGCTGTCAATCTACAGGACTTACTTCCATCCTGAGAG 360
QY 361 CTGGGCATCAATGCCAGCTTTTTCAGTCTTCCAAAGTCGGCTTAATACGATACCAGCTTT 420

DB 361 CTGGGCATCAATGCCAGCTTTTTCAGTCTTCCAAAGTCGGCTTAATACGATACCAGCTTT 420
QY 421 GTAGACAGGGGACTAGACGATCTCTGCACACTCTCTCCGGAGACAGAAGTATTGTTT 480
DB 421 GTAGACAGGGGACTAGACGATCTCTGCACACTCTCTCCGGAGACAGAAGTATTGTTT 480
QY 481 GATGTGTCCATCTCTCAGACAAAGAAGCTGTGGGGCGGAGCTGGGCTCTTTCCG 540
DB 481 GATGTGTCCATCTCTCAGACAAAGAAGCTGTGTGGGGCGGAGCTGGGCTCTTTCCG 540
QY 541 CAGGCGCCCTCAGCGCCCTGGGGCCACACGCGCGGCGCTCCAGCTGCGCTTCCCT 600
DB 541 CAGGCGCCCTCAGCGCCCTGGGGCCACACGCGCGGCGCTCCAGCTGCGCTTCCCT 600
QY 601 TGCCTTTGCCCTTACTGCTGACGCGGGAGCCCTGGACCCCGAGGGGCGCGCGGCC 660
DB 601 TGCCTTTGCCCTTACTGCTGACGCGGGAGCCCTGGACCCCGAGGGGCGCGCGGCC 660
QY 661 GGCTGGGAAGTCTTCGACGTGTGGCAGGCGCTGCCACACAGCCCTTGAAGCAGCTGTGC 720
DB 661 GGCTGGGAAGTCTTCGACGTGTGGCAGGCGCTGCCACACAGCCCTTGAAGCAGCTGTGC 720
QY 721 TTGGAGCTSCGGGCGCATTTGGGCGGAGCTGGACGCCGGGAGGCCGAGCGCGCGGG 780
DB 721 TTGGAGCTSCGGGCGCATTTGGGCGGAGCTGGACGCCGGGAGGCCGAGCGCGCGGG 780
QY 781 GGACCCCAACACCGCGCCCGCCGACCTTGGCGAGTCTGGGGCTTCGGCGGAGGTGCGG 840
DB 781 GGACCCCAACACCGCGCCCGCCGACCTTGGCGAGTCTGGGGCTTCGGCGGAGGTGCGG 840
QY 841 CTTCCCTCAGAGCGGCGCTTCTGTGTGTTTACACAGATCCACGCGAAGAACCTGTTC 900
DB 841 CTTCCCTCAGAGCGGCGCTTCTGTGTGTTTACACAGATCCACGCGAAGAACCTGTTC 900
QY 901 GCAGAGATGCGGAGCAGCTGGGCTCGGCGGAGGCTGGGGCGCGGGCGCGCGCGGAG 960
DB 901 GCAGAGATGCGGAGCAGCTGGGCTCGGCGGAGGCTGGGGCGCGGGCGCGCGCGGAG 960
QY 961 GGGTGTGGCGCGCGCTGCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGG 1020
DB 961 GGGTGTGGCGCGCGCTGCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGG 1020
QY 1021 CGCGCGCGCGCGCGCGCTTTCGCCAGTCCGCTTGGCAAGCGGACGCGAAGATCC 1080
DB 1021 CGCGCGCGCGCGCGCGCTTTCGCCAGTCCGCTTGGCAAGCGGACGCGAAGATCC 1080
QY 1081 AGGCTACGCTGCGAGCAAGAAGCCCTTGACGCTGAACCTCAAGGAGCTGGGCTGGGACGAC 1140
DB 1081 AGGCTACGCTGCGAGCAAGAAGCCCTTGACGCTGAACCTCAAGGAGCTGGGCTGGGACGAC 1140
QY 1141 TGGATTATCGCGCCCTGGAGTACGAGGCTATCCTGCGAGGGTGTATGCCACTTCCG 1200
DB 1141 TGGATTATCGCGCCCTGGAGTACGAGGCTATCCTGCGAGGGTGTATGCCACTTCCG 1200
QY 1201 CTGCGCTGCGACCTGGAGCCCAACCAACGCGCATCATCCAGACGCTGATGAATCCATG 1260
DB 1201 CTGCGCTGCGACCTGGAGCCCAACCAACGCGCATCATCCAGACGCTGATGAATCCATG 1260
QY 1261 GACCCCGGCTCCACCCCGCGCGAGCTGTGCTGCGCCACCAAAATTGACTCCCATACGATT 1320
DB 1261 GACCCCGGCTCCACCCCGCGCGAGCTGTGCTGCGCCACCAAAATTGACTCCCATACGATT 1320
QY 1321 CTATACATCGCGCGCAATAATGCTGTCTACAGCAGTACGAGGAGATGCTGTGGAG 1380
DB 1321 CTATACATCGCGCGCAATAATGCTGTCTACAGCAGTACGAGGAGATGCTGTGGAG 1380
QY 1381 TCGTGGGCTGCGAGTACGCTGCTTTCGCCGCGCGCTTGGCCG 1425
DB 1381 TCGTGGGCTGCGAGTACGCTGCTTTCGCCGCGCGCTTGGCCG 1425

RESULT 2

ABK48393
ID ABK48393 standard; DNA; 1425 BP.
AC
XX ABK48393;
XX
DT 02-JUL-2002 (first entry)
XX
DE DNA encoding human growth/differentiation factor 6-like protein NOV9.
XX
KW Human; growth/differentiation factor 6-like protein; NOV9; gene;
KW vaccine; cancer; tumour; bone disorder; avascular necrosis; allergy;
KW haematopoietic disorder; immune disorder; endometriosis; renal disease;
KW infection; inflammatory disease; lung disease; scleroderma; ataxia;
KW bowel disease; appendicitis; blood disorder; cardiovascular disorder;
KW graft versus host disease; GVHD; lymphoedema; brain disorder;
KW ocular disorder; hepatitis C virus infection; cardiac disorder;
KW autosomal dominant deafness; DFNA-2; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 7..1254
CDS /tag= a
ET /product= "Human growth/differentiation factor 6-like
ET protein NOV9"
XX
PN WO200214368-A2.
XX
XX 21-FEB-2002.
XX
PD 16-AUG-2001; 2001WO-US25624.
XX
PR 16-AUG-2000; 2000US-225692P.
PR 16-AUG-2000; 2000US-225693P.
PR 16-AUG-2000; 2000US-225837P.
PR 18-AUG-2000; 2000US-226236P.
PR 18-AUG-2000; 2000US-226353P.
PR 22-AUG-2000; 2000US-227085P.
PR 23-AUG-2000; 2000US-227395P.
PR 24-AUG-2000; 2000US-227492P.
PR 24-AUG-2000; 2000US-227600P.
PR 14-MAR-2001; 2001US-275952P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Zerhusen BD, Padigaru M, Spytek KA, Spaderna SK, Gangolli EA;
PI Rastelli L, Burgess CE, Majumder K, Shimkets R, Mishra V;
PI Vernet CAM, Szekeres ES, Grosse WM, Alsobrook JP, Liu X;
PI Gerlach VL, Ellerman K, Smithson G, Peyman J, Stone D;
PI MacDougall J;
XX
XX WPI; 2002-329571/36.
DR P-PSDB; AAU79173.
XX
XX Novel cytoplasmic, nuclear membrane bound and secreted NOVX
PT polypeptides, useful for treating cancers and tumours, bone disorders,
PT Paget's disease, haematopoietic disorders, spinal diseases and immune
PT disorders -
XX
PS Claim 8; Page 86; 234pp; English.
XX
XX The present invention relates to new isolated NOVX polypeptides named
CC NOV1-NOV9. The invention can be used for identifying an agent (a cellular
CC receptor or downstream effector) that binds to the polypeptide. The
CC molecules of the invention are useful for treating or preventing
CC NOVX-associated disorders in humans. The antibody of the invention is
CC useful for determining the presence or amount of NOVX in a sample, and
CC for treating a pathological state in a mammal. The method of the
CC invention is useful for determining the presence of an amount of NOVX in
CC a sample which is used as a marker for cancerous cell or tissue type.
CC The molecules of the invention are useful in the manufacture of a
CC medicament for treating or preventing cancer, tumour, bone disorders,
CC avascular necrosis, allergy, haematopoietic disorders, immune disorders,

CC endometriosis, renal diseases, infections, inflammatory diseases, lung
CC diseases, scleroderma, ataxia, bowel diseases, appendicitis, blood
CC disorders, cardiovascular disorders, graft versus host disease (GVHD),
CC lymphoedema, brain disorders, ocular disorders, hepatitis C virus
CC infection, cardiac disorders and autosomal dominant deafness (DFNA-2).
CC The present nucleic acid sequence encodes the human
CC growth/differentiation factor 6-like protein NOV9 that is one of the
CC NOVX proteins described in the invention.
XX
SQ Sequence 1425 BP; 234 A; 488 C; 459 G; 234 T; 0 other;
XX
Query Match 100.0%; Score 1425; DB 24; Length 1425;
Best Local Similarity 100.0%; Pred. No. 6.7e-246;
Matches 1425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTGGGAGAGCGACGACCTTGCCTGATGATCTCCAGGCTTCCATCTCATCC 60
DB 1 CTCCTGGGAGAGCGACGACCTTGCCTGATGATCTCCAGGCTTCCATCTCATCC 60
QY 61 GTCCTCCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120
DB 61 GTCCTCCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGCAGGCTTCCATCTCATCC 120
QY 121 TCTGTGTTCTGCGCGAGCTGGGTTCCACCAAGGGCATCGAAGCGCGAAGGCAAG 180
DB 121 TCTGTGTTCTGCGCGAGCTGGGTTCCACCAAGGGCATCGAAGCGCGAAGGCAAG 180
QY 181 ATGCAGCGGGCGCGCGGACAGTACGCGCGGGGCGGAGGCGCAGCAACGCGCGG 240
DB 181 ATGCAGCGGGCGCGCGGACAGTACGCGCGGGGCGGAGGCGCAGCAACGCGCGG 240
QY 241 CCTCAGGAGCAACCCCGGGCTCAGCAGCGCCCGGCGCAGGCGCGGAGGGTCCG 300
DB 241 CCTCAGGAGCAACCCCGGGCTCAGCAGCGCCCGGCGCAGGCGCGGAGGGTCCG 300
QY 301 CGCGTGGTCCCGCCAGAGTACATGCTGTCATCTACAGGACTTACTTCCATCGTGAGAAG 360
DB 301 CGCGTGGTCCCGCCAGAGTACATGCTGTCATCTACAGGACTTACTTCCATCGTGAGAAG 360
QY 361 CTGGGATCAATGCCAGCTTTTCCAGTCTTCCAGTGGCTTCCAGTGGCTTCCAGTGGCTT 420
DB 361 CTGGGATCAATGCCAGCTTTTCCAGTCTTCCAGTGGCTTCCAGTGGCTTCCAGTGGCTT 420
QY 421 GTAGCAGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTT 480
DB 421 GTAGCAGGGGACTAGACGATCTCTCGCACACTCTCTCCGGAGACAGAGTATTTGTTT 480
QY 481 GATGTGTCATGCTCTCAGACAAAGAGCTGTGGGGCGGAGGCTGCGGCTCTTTTCGC 540
DB 481 GATGTGTCATGCTCTCAGACAAAGAGCTGTGGGGCGGAGGCTGCGGCTCTTTTCGC 540
QY 541 CAGGCGCCCTCAGGCGCCCTGGGGCCACAGCGGGCGGCTCCACGTGAGCTCTTCCCT 600
DB 541 CAGGCGCCCTCAGGCGCCCTGGGGCCACAGCGGGCGGCTCCACGTGAGCTCTTCCCT 600
QY 601 TGCCTTTCGCCCTTACTGCTGGACGCGGAGCCCTGACCCCGCAGGGGCGCGCGCGCC 660
DB 601 TGCCTTTCGCCCTTACTGCTGGACGCGGAGCCCTGACCCCGCAGGGGCGCGCGCGCC 660
QY 661 GCGTGGGAAGTCTTCGACGTGTGGAGGGCTTGGCCACACAGCCCTGGGAAGCAGCTGTGC 720
DB 661 GCGTGGGAAGTCTTCGACGTGTGGAGGGCTTGGCCACACAGCCCTGGGAAGCAGCTGTGC 720
QY 721 TTGGAGCTCGGGCGCATGGGCGAGCTGGACGCGGGGAGGCGGAGCGCGCGCGG 780
DB 721 TTGGAGCTCGGGCGCATGGGCGAGCTGGAGCGCGGGGAGGCGGAGCGCGCGCGG 780
QY 781 GGACCCCAACACCGCGCCCGCCGACCTGCGGAGTCTGGGCTTGGCGGAGGGTGC 840
DB 781 GGACCCCAACACCGCGCCCGCCGACCTGCGGAGTCTGGGCTTGGCGGAGGGTGC 840
QY 841 CCTCCCGCAGGAGCGGCGCTCTGTGTGTTTACACAGATCCACGCGCAAGACCTGTC 900
DB 841 CCTCCCGCAGGAGCGGCGCTCTGTGTGTTTACACAGATCCACGCGCAAGACCTGTC 900

Db 841 CCTCCCCAGGAGCGGGCCCTGCTGTTGTTATTCACAGATCCAGCGCAAGAACCTGTTTC 900
QY 901 GCAGAGATCGCGAGCAGCTGGCTCGGCCGAGAGGTGCGGGCCCGCGCGCGCCGAG 960
Db 901 GCAGAGATCGCGAGCAGCTGGCTCGGCCGAGAGGTGCGGGCCCGCGCGCGCCGAG 960
QY 961 GGGTCTGTGCGCCCGCGCTGCGGGCGCCCGGATGCGCAGGCTTGGCTGCGCCCTGCGCCGCGC 1020
Db 961 GGGTCTGTGCGCCCGCGCTGCGGGCGCCCGGATGCGCAGGCTTGGCTGCGCCCTGCGCCGCGC 1020
QY 1021 CGCCGGCGGGCGACAGCGCTTCGGCAGTGCAGCGGATGGCAAGCGGACAGGCAAGATGCC 1080
Db 1021 CGCCGGCGGGCGACAGCGCTTCGGCAGTGCAGCGGATGGCAAGCGGACAGGCAAGATGCC 1080
QY 1081 AGGCTACGCTGCAGCAAGAACGCCCTGACGCTGAACCTTCAAGGAGCTGGGCTGGGACGAC 1140
Db 1081 AGGCTACGCTGCAGCAAGAACGCCCTGACGCTGAACCTTCAAGGAGCTGGGCTGGGACGAC 1140
QY 1141 TGGATTATCGCGCCCTCGAGTACGAGGCTATCACTGCGAGGGTGTATGCCACTTCCCG 1200
Db 1141 TGGATTATCGCGCCCTCGAGTACGAGGCTATCACTGCGAGGGTGTATGCCACTTCCCG 1200
QY 1201 CTGCGCTCGCAGCTGGAGCGCCACCAACACGCCATCATCCAGACGCTGATGAACCTCCATG 1260
Db 1201 CTGCGCTCGCAGCTGGAGCGCCACCAACACGCCATCATCCAGACGCTGATGAACCTCCATG 1260
QY 1261 GACCCGGCTCCACCCCGCCAGCTGCTGCGTGGCCACCAAAATGACTCCCATCAGCAT 1320
Db 1261 GACCCGGCTCCACCCCGCCAGCTGCTGCGTGGCCACCAAAATGACTCCCATCAGCAT 1320
QY 1321 CTATACATCGACGCGGCAATATGTGTTCTACAAGCAGTACGAGGACATGTTGGTGGAG 1380
Db 1321 CTATACATCGACGCGGCAATATGTGTTCTACAAGCAGTACGAGGACATGTTGGTGGAG 1380
QY 1381 TCGTGGGCTCGAGTAGCGGTGCTTTCCCGCGCGCTTGGCCCG 1425
Db 1381 TCGTGGGCTCGAGTAGCGGTGCTTTCCCGCGCGCTTGGCCCG 1425

RESULT 3
AAD28496
ID AAD28496 standard; cDNA; 2791 BP.
XX AAD28496;
DT
XT
XX 22-APR-2002 (first entry)
DE Human extracellular messenger (XMES)-6 cDNA.
XX Human; extracellular messenger; neurological disorder; epilepsy; XMES-6;
KW Alzheimer's disease; autoimmune disorder; renal tubular acidosis; stroke;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; cytostatic;
KW developmental disorder; anaemia; Cushing's syndrome; endocrine disorder;
KW vascular malformation; cell proliferative disorder; gene therapy; cancer;
KW neuroprotective; trauma; hypopituitarism; hypothyroidism; antihelminthic;
KW hyperthyroidism; gonadal steroid hormone; pancreatic disorder; nontropic;
KW diabetes mellitus; immunosuppressive; anti-inflammatory; antibacterial;
KW antiviral; antifungal; parasitic; protozoal; allergy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1029..2396
CDS /*tag= a
FT /*product= "Human XMES-6 protein"
FT
XX WO200194587-A2.
PN
XX
XX 13-DEC-2001.
XX
XX 06-JUN-2001; 2001WO-US18476.
PF
XX
XX 06-JUN-2000; 2000US-210233P.
PR

PR 23-JUN-2000; 2000US-213465P.
PR 14-NOV-2000; 2000US-249019P.
XX (INCY-) INCYTE GENOMICS INC.
PA
PI Lal P, Yue H, He A, Nguyen DB, Wallia N, Gandhi AR, Azimzai Y;
PI Bandman O, Tang YT, Lu Y, Baughn MR, Duggan BM, Lee S, Hafalia A;
PI Policky JL;
XX
DR WPI; 2002-154573/20.
DR P-PSDB; AAE17604.
XX
PT New extracellular messenger polypeptides and polynucleotides encoding
PT them, useful for diagnosing, treating or preventing e.g. neurological,
PT autoimmune, inflammatory, developmental and endocrine disorders -
XX
PS Claim 5; Page 120-121; 123pp; English.
XX
CC The present invention relates to new extracellular messenger polypeptides
CC and polynucleotides encoding them. XMES is useful in the diagnosis,
CC treatment and prevention of neurological disorders (e.g. epilepsy,
CC stroke, or Alzheimer's disease), autoimmune/inflammatory disorders (e.g.
CC acquired immune deficiency syndrome, AIDS, Addison's disease, or
CC allergies), developmental disorders (e.g. renal tubular acidosis, anaemia
CC or Cushing's syndrome), endocrine disorders (e.g. hypophysectomy,
CC aneurysm or vascular malformation), and cell proliferative disorders
CC (e.g. cancer), and in the assessment of the effects of exogenous
CC compounds on the expression of nucleic acid and amino acid sequences of
CC extracellular messengers. XMES may also be used in the treatment of
CC viral, bacterial, fungal, parasitic, protozoal and helminthic
CC infections, trauma, disorders associated with hypopituitarism,
CC hypothyroidism, hyperthyroidism or gonadal steroid hormones, and
CC pancreatic disorders such as type I or type II diabetes mellitus. The
CC XMES may be used for somatic or germline gene therapy. The present
CC sequence is human XMES-6 cDNA.
XX
SQ Sequence 2791 BP; 454 A; 945 C; 942 G; 450 T; 0 other;
Query Match 99.8%; Score 1421.8; DB 24; Length 2791;
Best Local Similarity 99.9%; Pred. No. 2.5e-245;
Matches 1423; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CTCCTGGGAGACGACGACCTTGCCCGCCATGGATATCCAGGGTCTGCTCTCGGCC 60
Db 999 CTCCTGGGAGACGACGACCTTGCCCGCCATGGATATCCAGGGTCTGCTCTCGGCC 1058
QY 61 GTCCTCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGAGGCTTCCATCTCATCC 120
Db 1059 GTCCTCTCATCAGTTTCTGTGGGATTTGCCGGTTTCCAGAGGCTTCCATCTCATCC 1118
QY 121 TCCTGTTCTCGCCGAGCTGGTTTCCACCAAGGSCATGCCAAGCGCAAGAGGCAAG 180
Db 1119 TCCTGTTCTCGCCGAGCTGGTTTCCACCAAGGSCATGCCAAGCGCAAGAGGCAAG 1178
QY 181 ATGACGCGGGCGCGCGACAGTGTACGCGGGCGGGAGGCGCCAGAACACAGCGCGG 240
Db 1179 ATGACGCGGGCGCGCGACAGTGTACGCGGGCGGGAGGCGCCAGAACACAGCGCGG 1238
QY 241 CCTAGGAGCAACCCCGGGCTCAGAGCCCCGGGGAGGAGCCCGCAGGAGGGTCCG 300
Db 1239 CCTAGGAGCAACCCCGGGCTCAGAGCCCCGGGGAGGAGCCCGCAGGAGGGTCCG 1298
QY 301 CGCGTGGTGGCCCGAGGATACATGCTGCAATCTACAGACTTACTCCATCGCTGAGAAG 360
Db 1299 CGCGTGGTGGCCCGAGGATACATGCTGCAATCTACAGACTTACTCCATCGCTGAGAAG 1358
QY 361 CTGGGCATCAATGCGCAGCTTTTTCAGTCTTCCAAGTCCGCTACATACGATCACCAGCTTT 420
Db 1359 CTGGGCATCAATGCGCAGCTTTTTCAGTCTTCCAAGTCCGCTACATACGATCACCAGCTTT 1418
QY 421 GTAGACGGGGACTAGAGATCTCTCGCACACTCTCTCCGGAGACAGAAATTTGTTT 480
Db 1419 GTAGACGGGGACTAGAGATCTCTCGCACACTCTCTCCGGAGACAGAAATTTGTTT 1478


```
QY 481 GATGTGTCCATGCTCTCAGACAAAGAGAGCTGGTGGGCGGAGCTGGGCTCTTTTCGC 540
   |||||||
Db 1479 GATGTGTCCATGCTCTCAGACAAAGAGAGCTGGTGGGCGGAGCTGGGCTCTTTTCGC 1538
QY 541 CAGGCGCCCTCAGCGCCCTGGGGGCGACACAGCGGGCGGCTCCACGTGAGCTCTTCCCT 600
   |||||||
Db 1539 CAGGCGCCCTCAGCGCCCTGGGGGCGACACAGCGGGCGGCTCCACGTGAGCTCTTCCCT 1598
QY 601 TGGCTTTGGCCCTACTGCTGGACGCGGACCTCGACCCGAGGGGCGCGCGCGG 660
   |||||||
Db 1599 TGGCTTTGGCCCTACTGCTGGACGCGGACCTCGACCCGAGGGGCGCGCGCGG 1658
QY 661 GGTGGGAAGTCTTCGACGCTGTGGCAGGGCTCGCGCACAGCCCTGGAAGCAGCTGTGC 720
   |||||||
Db 1659 GGTGGGAAGTCTTCGACGCTGTGGCAGGGCTCGCGCACAGCCCTGGAAGCAGCTGTGC 1718
QY 721 TTGGAGCTGGGGCGCGATGGGGCGGAGCTGGAGCCGGGGAGGCCAGGGCGCGCGG 780
   |||||||
Db 1719 TTGGAGCTGGGGCGCGATGGGGCGGAGCTGGAGCCGGGGAGGCCAGGGCGCGCGG 1778
QY 781 GGACCCAGCAACCGCGCCCGGACCTGCGGAGTCTGGGCTTCGCGGAGGGTGCGG 840
   |||||||
Db 1779 GGACCCAGCAACCGCGCCCGGACCTGCGGAGTCTGGGCTTCGCGGAGGGTGCGG 1838
QY 841 CTTCCCCAGGAGCGGCGCTGCTGGTGGTATTCACAGATCCCGAGCGCAAGAACCTGTTC 900
   |||||||
Db 1839 CTTCCCCAGGAGCGGCGCTGCTGGTGGTATTCACAGATCCCGAGCGCAAGAACCTGTTC 1898
QY 901 GCAGATGCGCGAGCAGCTGGGCTCGCGGAGGCTGGCGGCGCGGCGCGCGCGAG 960
   |||||||
Db 1899 GCAGATGCGCGAGCAGCTGGGCTCGCGGAGGCTGGCGGCGCGGCGCGCGCGAG 1958
QY 961 GGGTCGTGGCGCGCGCTGCGGCGCCCGGATGCCAGGCTTGGCTGCGCCCGCGG 1020
   |||||||
Db 1959 GGGTCGTGGCGCGCGCTGCGGCGCCCGGATGCCAGGCTTGGCTGCGCCCGCGG 2018
QY 1021 CGCGGCGGCGCGCACGGGCTTCGCGAGTTCGCGAGTGGCAAGCGGCACGCAAGATCC 1080
   |||||||
Db 2019 CGCGGCGGCGCGCACGGGCTTCGCGAGTGGCAAGCGGCACGCAAGATCC 2078
QY 1081 AGGCTAGCTGCAGCAAGAACGCCCTGCGAGTGAACTTCAAGAGCTGGGCTGGGACGAC 1140
   |||||||
Db 2079 AGGCTAGCTGCAGCAAGAACGCCCTGCGAGTGAACTTCAAGAGCTGGGCTGGGACGAC 2138
QY 1141 TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGGGAGGCTGTATGCGACTTCCCG 1200
   |||||||
Db 2139 TGGATTATCGCGCCCTGGAGTACGAGGCTATCACTGGGAGGCTGTATGCGACTTCCCG 2198
QY 1201 CTGCGTGGCACCTGGAGGCCCAACCAAGCCATCATCCAGACGCTGATGAACCTCCATG 1260
   |||||||
Db 2199 CTGCGTGGCACCTGGAGGCCCAACCAAGCCATCATCCAGACGCTGATGAACCTCCATG 2258
QY 1261 GACCCGGCTCCACCGCCGCGGCTGCTGGTGGTCCCAACCAATGACTCCCATCAGATT 1320
   |||||||
Db 2259 GACCCGGCTCCACCGCCGCGGCTGCTGGTGGTCCCAACCAATGACTCCCATCAGATT 2318
QY 1321 CTATACATGACCGCGCAATATGTTCTACAACAGTACGAGGACATGGTGGAG 1380
   |||||||
Db 2319 CTATACATGACCGCGCAATATGTTCTACAACAGTACGAGGACATGGTGGAG 2378
QY 1381 TCGTGGGCTGCAAGGTAGCGGTGCTTTCCCGCGGCTTTGGCCCG 1425
   |||||||
Db 2379 TCGTGGGCTGCAAGGTAGCGGTGCTTTCCCGCGGCTTTGGCCCG 2423
```

RESULT 4

AAT90386
ID AAT90386 standard; DNA; 1171 BP.
XX
AC AAT90386;

XX 21-JAN-1998 (first entry)

```
XX Human bone morphogenetic protein BMP-13 gene clone vl-1.
DE BMP-13; bone morphogenetic protein; human; tendon; ligament;
KW wound healing; tissue repair; tendonitis; carpal tunnel syndrome;
KW therapy; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 2..967
FT /*tag= a
FT sig_peptide 2..604
FT /*tag= b
FT mat_peptide 605..964
FT /*tag= c
XX US5658882-A.
PN 19-AUG-1997.
XX 07-DEC-1993; 93US-0164103.
XX 22-DEC-1994; 94US-0362670.
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX (GEMY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
PI Wozney JM;
XX WPI; 1997-424270/39.
DR P-PSDB; AAW26591.
XX Inducing tendon and ligament formation using BMP-12, BMP-13 or MP-52
PT - useful for tissue healing and repair, treatment of tendonitis,
PT improving fixation of tendons to bone etc
XX Example 1; Column 55-58; 43pp; English.
XX This nucleotide sequence, designated vl-1, encodes a novel bone
CC morphogenetic protein (see AAW26591), designated BMP-13. It was
CC isolated from a human genomic library using a probe (see AAT90398)
CC based on a partial BMP-13 sequence. A claimed method for inducing
CC formation of tendon and/or ligament tissues involves administration
CC of a composition containing at least one protein selected from
CC BMP-13, MP52 (see AAW26590) and BMP-12 (see AAW26589). This is
CC used for tissue (including skin) healing and repair. This is
CC useful for treating tendonitis, carpal tunnel syndrome and other
CC defects of traumatic or congenital origin, in cosmetic surgery and
CC to improve fixation of tendons or ligaments to bone. The specified
CC proteins can also be used to increase activity of other BMPs e.g.
CC BMP-2 (see AAW26597).
XX Sequence 1171 BP; 195 A; 384 C; 413 G; 179 T; 0 other;
SQ Query Match 69.3%; Score 987; DB 18; Length 1171;
Best Local Similarity 100.0%; Pred. NO. 1.2e-167;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGATTTGTTTGATGTGTCATGCTCTCA 498
|||||
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGATTTGTTTGATGTGTCATGCTCTCA 67

QY 499 GACAAAGAGAGCTGGTGGCGCGGAGCTGCGGCTCTTTTCGCGAGCGCCCTCAGCGCCC 558
|||||
Db 68 GACAAAGAGAGCTGGTGGCGCGGAGCTGCGGCTCTTTTCGCGAGCGCCCTCAGCGCCC 127

QY 559 TGGGGGCCACAGCGCGGCGGCTCCACGCTTCCTTGGCTTCCCTTGGCCCTTACTG 618
|||||

```
Db 128 TGGGGCCACACAGCGGGCGCTCCACGTGACGTCTTCCCTTGCCCTTTCGCCCTACTG 187
QY 619 CTGGAGCGGGACCTTGACCGCGAGGGGGCGCCGCCGGCTGSGAAGTCTTCGAC 678
Db 188 CTGGAGCGGGACCTTGACCGCGAGGGGGCGCCGCCGGCTGSGAAGTCTTCGAC 247
QY 679 GTGTGGCAGGGCCCTGGCCACACAGCCCTTGAAGCAGCTGTGCTTGAGCTGGGGCGCA 738
Db 248 GTGTGGCAGGGCCCTGGCCACACAGCCCTTGAAGCAGCTGTGCTTGAGCTGGGGCGCA 307
QY 739 TGGGGCAGCTGAGCGCGGGGAGGCGCGCGCGGGGACCCAGCAACCGCGG 798
Db 308 TGGGGCAGCTGAGCGCGGGGAGGCGCGCGCGGGGACCCAGCAACCGCGG 367
QY 799 CCCCCGAGCTGGGAGTCTGGGCTTCGGCCGAGGGTGGGGCTCCCGAGGAGCGGGC 858
Db 368 CCCCCGAGCTGGGAGTCTGGGCTTCGGCCGAGGGTGGGGCTCCCGAGGAGCGGGC 427
QY 859 CTGCTGGTGTATTACACAGATCCACAGCGCAAGAACTGTTCGACAGATGCGCAGCAG 918
Db 428 CTGCTGGTGTATTACACAGATCCACAGCGCAAGAACTGTTCGACAGATGCGCAGCAG 487
QY 919 CTGGGCTCGGCGAGGCTCGGGGCCCGGGCGGGCGCGAGGGTCTGTGGCGCGCGG 978
Db 488 CTGGGCTCGGCGAGGCTCGGGGCCCGGGCGGGCGCGAGGGTCTGTGGCGCGCGG 547
QY 979 TCGGGCGCCCGGATGCCAGGCTTGGCTGCGCTCGCCGCGCGCGCGGGCGGCGCAG 1038
Db 548 TCGGGCGCCCGGATGCCAGGCTTGGCTGCGCTCGCCGCGCGCGGGCGGCGCAG 607
QY 1039 GCCTTCGCCAGTGCCTATGCAAGCGGCACGGCAAGATGCCAGGCTACGCTGCAGCAAG 1098
Db 608 GCCTTCGCCAGTGCCTATGCAAGCGGCACGGCAAGATGCCAGGCTACGCTGCAGCAAG 667
QY 1099 AAGCCCTCGACGTGAACCTCAAGGAGCTGGGCTGGGAGCTGATATTCGGCCCTCG 1158
Db 668 AAGCCCTCGACGTGAACCTCAAGGAGCTGGGCTGGGAGCTGATATTCGGCCCTCG 727
QY 1159 GAGTACGAGGCTTACATCGCAGGGGTATGCGACTTCCCGCTGCGCTGCACCTGGAG 1218
Db 728 GAGTACGAGGCTTACATCGCAGGGGTATGCGACTTCCCGCTGCGCTGCACCTGGAG 787
QY 1219 CCACCAACACGCCATCATCCAGACGCTGATGAATCCATGACGCCGCCGCTCCACCCG 1278
Db 788 CCACCAACACGCCATCATCCAGACGCTGATGAATCCATGACGCCGCCGCTCCACCCG 847
QY 1279 CCAGTGTGCTGGTGGCCACCAAAATGATCCCATCAGCATTCATATACATCAGCGGGC 1338
Db 848 CCAGTGTGCTGGTGGCCACCAAAATGATCCCATCAGCATTCATATACATCAGCGGGC 907
QY 1339 AATAATGTGGTCTACAGCAGTACGAGGACATGGTGTGAGTCTGCGGCTGCAGGTAG 1398
Db 908 AATAATGTGGTCTACAGCAGTACGAGGACATGGTGTGAGTCTGCGGCTGCAGGTAG 967
QY 1399 CGGTGCTTTCCCGCGCCCTTGCGCCG 1425
Db 968 CGGTGCTTTCCCGCGCCCTTGCGCCG 994
```

RESULT 5

```
AAD18333
ID AAD18333 standard; DNA; 1171 BP.
XX
AC AAD18333;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human full length VL-1 or BMP-13 DNA.
XX
KW Human; bone morphogenic protein; BMP-13; vulnery; antiinflammatory;
KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;
KW tissue formation; wound healing; tissue repair; ligament defect; VL-1;
KW carpal tunnel syndrome; tendonitis; ds.
```

```
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 2..967
FT /tag= a
FT /product= "Human VL-1 protein"
FT /note= "CDS does not include start codon"
FT /partial
FT sig_peptide 2..604
FT mat_peptide 605..964
FT /tag= b
FT /tag= c
FT /product= "Mature VL-1 protein"
XX
PN US284872-B1.
XX
PD 04-SEP-2001.
XX
PF 28-FEB-1997; 97US-0808324.
XX
PR 22-DEC-1994; 94US-0362670.
PR 07-DEC-1993; 93US-0164103.
PR 02-MAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX
PA (GENY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
PI Celeste AJ, Wozney JM, Rosen VA, Wolfin NM, Thomsen GH;
PI Melton DA;
XX
XX WPI: 2001-588978/66.
DR P-PSDB; AAE10982.
XX
PT New chimeric DNAs, useful for treating tendonitis, carpal tunnel
PT syndrome and other tendon and ligament defects, comprises DNA encoding
PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
PT BMP-13 or MP52 -
XX
XX Example 1; Column 55-58; 42pp; English.
XX
CC The invention relates to a chimeric DNA comprising a DNA sequence
CC encoding a propeptide from a member of the transforming growth factor
CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
CC encoding an amino acid sequence encoding a mature polypeptide consisting
CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
CC sequences are useful for producing proteins which induce tendon/ligament
CC like tissue formation, and for isolating and cloning further DNA
CC sequences encoding BMP-12 related proteins with similar activity. The
CC proteins are useful for the induction of tendon/ligament-like tissue
CC formation, wound healing, ligament and other tissue repair, augmenting
CC the activity of bone morphogenetic proteins, and for treating tendonitis,
CC carpal tunnel syndrome and other tendon and ligament defects. The
CC present sequence is human VL-1 DNA fragment also designated as
CC BMP-13.
XX
SQ Sequence 1171 BP; 195 A; 384 C; 413 G; 179 T; 0 other;
Query Match 69.3%; Score 987; DB 22; Length 1171;
Best Local Similarity 100.0%; Pred. No. 1.2e-167;
Matches 987; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGCTCA 498
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGTATTGTTGATGTGCTCA 67
QY 499 GACAAAGAGAGCTGGTGGCGGAGCTGGCGCTTTTCGCCAGCGCCCTCAGCGCC 558
Db 68 GACAAAGAGAGCTGGTGGCGGAGCTGGCGCTTTTCGCCAGCGCCCTCAGCGCC 127
QY 559 TGGGGGCCACACAGCGCGCGCTCCACGTGACGTCTTCCCTTGCCTTTCGCCCTACTG 618
|||||
```

Db 128 TGGGGGCCACACCGCGCGCTCCAGTGCAGCTTTCCTTGGCTTTCGCCCTTACTG 187
QY 619 CTGGACGCGCGGACCTTGGACCCGCGAGGGGCGCCCGCGCGCTGGGAAGTCTTCGAC 678
Db 188 CTGGACGCGCGGACCTTGGACCCGCGAGGGGCGCGCGCGCTGGGAAGTCTTCGAC 247
QY 679 GTGTGCGAGGGCTGCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGTGGGGCGCA 738
Db 248 GTGTGCGAGGGCTGCGCCACACAGCCCTTGAAGCAGCTGTGCTTGGAGTGGGGCGCA 307
QY 739 TGGGCGAGCTGACCGCGGGAGGCGGAGCGCGCGCGCGGAGCCCGAGCAACCGCG 798
Db 308 TGGGCGAGCTGACCGCGGGAGGCGGAGCGCGCGCGGAGCCCGAGCAACCGCG 367
QY 799 CCCCCGACCTGCGGAGTCTGGGCTTTCGCGCGGAGGCTGCGGCTCCCGAGAGCGGGC 858
Db 368 CCCCCGACCTGCGGAGTCTGGGCTTTCGCGCGGAGGCTGCGGCTCCCGAGAGCGGGC 427
QY 859 CTGCTGGTGTATTCACAGATCCAGCGGAGAGAACCTGTTCGACAGATGCGCGAGCAG 918
Db 428 CTGCTGGTGTATTCACAGATCCAGCGGAGAGAACCTGTTCGACAGATGCGCGAGCAG 487
QY 919 CTGGGCTGCGGCGAGCTGCGGCGCGGCGCGGCGCGGCGGCGGCGGCGCGCG 978
Db 488 CTGGGCTGCGGCGAGCTGCGGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCG 547
QY 979 TCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCG 1038
Db 548 TCGGGCGCGCGGATGCCAGGCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCG 607
QY 1039 GCCTTGCAGTGCATGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1098
Db 608 GCCTTGCAGTGCATGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 667
QY 1099 AAGCCCTGCGAGTGAATTCAGGAGCTGCGGCTGGGAGGCTGATTCGCGCGCGCTG 1158
Db 668 AAGCCCTGCGAGTGAATTCAGGAGCTGCGGCTGGGAGGCTGATTCGCGCGCGCTG 727
QY 1159 GAGTAGAGGCTTATCAGTCCGAGGCTGTATGCGACTTCCCGCTCGGAGCTGGAG 1218
Db 728 GAGTAGAGGCTTATCAGTCCGAGGCTGTATGCGACTTCCCGCTCGGAGCTGGAG 787
QY 1219 CCACCAACGACCCATATCAGAGCGCTGATGAATCCATGAGACCCCGGCTCCACCCCG 1278
Db 788 CCACCAACGACCCATATCAGAGCGCTGATGAATCCATGAGACCCCGGCTCCACCCCG 847
QY 1279 CCAGTGTGCTGCGTCCGACCAATTTGACTCCCATCAGCATCTATATCAGCGGCG 1338
Db 848 CCAGTGTGCTGCGTCCGACCAATTTGACTCCCATCAGCATCTATATCAGCGGCG 907
QY 1339 AATAATGTGCTTACAGCAGTACGAGGACATGGTGTGAGTCTGCGGCTGCGAGTAG 1398
Db 908 AATAATGTGCTTACAGCAGTACGAGGACATGGTGTGAGTCTGCGGCTGCGAGTAG 967
QY 1399 CGGTGCTTTCGCGCGGCTTGGCCCG 1425
Db 968 CGGTGCTTTCGCGCGGCTTGGCCCG 994

RESULT 6

AAQ96208
ID AAQ96208 standard; DNA; 1171 BP.
XX AC AAQ96208;
XX 30-NOV-1995 (first entry)
XX Human mature VL-1 (BMP-13) encoding sequence.
XX Bone morphogenetic protein; VL-1; tendon; ligament; tendinitis; ss.
XX Homo sapiens.

FH Key Location/Qualifiers
FT CDS 2..964
FT /tag= a
FT mat_peptide 605..964
FT /tag= b
FT misc_difference 671..874
FT /tag= c
FT /note= "degenerate codes used due to poor print quality"
FT misc_feature 605..964
FT /tag= d
FT /note= "claimed"
FT misc_feature 659..964
FT /tag= e
FT /note= "claimed"
PN W09516035-A.
XX 15-JUN-1995.
XX 06-DEC-1994; 94WO-US14030.
XX 02-NOV-1994; 94US-0333576.
XX 07-DEC-1993; 93US-0164103.
XX 25-MAR-1994; 94US-0217780.
XX (GEMV) GENETICS INST INC.
XX (HARD) HARVARD COLLEGE.
PI Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
PI Wozney JM;
XX P-PSDB; AAR78730.
DR WPI: 1995-224320/29.
XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in compsn. for inducing tendon/ligament-like tissue formation
PT Claim 2; Page 62-64; 84pp; English.
XX BMP-12 related proteins are a subset of the BMP/TGF-beta/Vg-1 family of proteins, including BMP-12 and VL-1. VL-1 is designated BMP-13. Like BMP-12, it is expected that BMP-13, as expressed by mammalian cells such as CHO cells, exists as a heterogeneous popn. of active species of BMP-13 protein with varying N-termini. It is expected that all active species will contain the AA sequence beginning with the 19th Cys residue of the mature protein until the 119th residue of the mature protein or until the stop codon after the 120th residue of the mature protein. Other active species contain additional AA sequence in the N-terminal direction. CC AAQ96208 is a partial DNA sequence and AAR78730 is the derived AA sequence of a portion of the 2.5 kb DNA insert of the plasmid subclone pGEMJIDc31/2.5, derived from clone lambdaJIDc31.
SQ Sequence 1171 BP; 195 A; 377 C; 408 G; 174 T; 4 U; 13 other;

Query Match 68.5%; Score 976; DB 16; Length 1171;
Best Local Similarity 98.2%; Pred. No. 1,1e-165;
Matches 969; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 439 GATCTCTCGCACACTCTCTCCGGAGACAGAGAATTTTGTGATGTGCTTCA 498
Db 8 GATCTCTCGCACACTCTCTCCGGAGACAGAGAATTTTGTGATGTGCTTCA 67
QY 499 GACAAAGAGAGCTGTGGGCGGAGCTTTCGCCAGCGCCCTCAGGCGCC 558
Db 68 GACAAAGAGAGCTGTGGGCGGAGCTTTCGCCAGCGCCCTCAGGCGCC 127
QY 559 TGGGGGCCACACCGCGCGCTCCACGTGCAGCTTCCCTTGCCTTTCGCCCTTACTG 618
Db 128 TGGGGGCCACACCGCGCGCTCCACGTGCAGCTTCCCTTGCCTTTCGCCCTTACTG 187
QY 619 CTGGACGCGCGGACCTGGACCGCGGCGCGCGCGGCTGGGAAGTCTTCGAC 678

```
Db 188 CTGAGCGCGGACCCCTGGACCGCGAGGGGCGCCGCGCGCTGGGAAGTCTTCGAC 247
Qy 679 GTGTGGCAGGCTCGCGCCACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGGGGCCGCA 738
Db 248 GTGTGGCAGGCTCGCGCCACAGCCCTGGAAGCAGCTGTGCTTGGAGCTGGGGCCGCA 307
Qy 739 TGGGGCAGCTGGACCGCGGGAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 798
Db 308 TGGGGCAGCTGGACCGCGGGAGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 367
Qy 799 CCCCCGAGCTGGGAGCTGTGGCTTTCGCGCGGAGGCTCGCGCGCTCCCGAGGCGGCC 858
Db 368 CCCCCGAGCTGGGAGCTGTGGCTTTCGCGCGGAGGCTCGCGCGCTCCCGAGGCGGCC 427
Qy 859 CTGCTGGTGTATTCACAGATCCACAGCCAGAACCTGTTCGACAGATGCGCGAGCAG 918
Db 428 CTGCTGGTGTATTCACAGATCCACAGCCAGAACCTGTTCGACAGATGCGCGAGCAG 487
Qy 919 CTGGGCTCGCGGAGGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 978
Db 488 CTGGGCTCGCGGAGGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 547
Qy 979 TGGGGCGCCCCGAGTCCAGGCTTGGCTGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 1038
Db 548 TGGGGCGCCCCGAGTCCAGGCTTGGCTGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 607
Qy 1039 GCTTCGCGAGTCCGATCGCAGCGGCGCAGCGCAGAACCTCCAGGCTACGCTGCAGCAAG 1098
Db 608 GCTTCGCGAGTCCGATCGCAGCGGCGCAGCGCAGAACCTCCAGGCTACGCTGCAGCAAG 667
Qy 1099 AAGCCCTCGACGTGAATCAAGAGCTGGGCTGGGAGCTGGATTATCGCGCCCTG 1158
Db 668 AAGCCCTCGACGTGAATCAAGAGCTGGGCTGGGAGCTGGATTATCGCGCCCTG 727
Qy 1159 GAGTAGGCGCTATCAGTCGAGGCTGTATGGAGTTCGCGCTGGCTGGCTGGCTGGAG 1218
Db 728 GAGTAGGCGCTATCAGTCGAGGCTGTATGGAGTTCGCGCTGGCTGGCTGGAG 787
Qy 1219 CCACCAACCGCCATCATCAGAGCTGTATGAGTCCATGAGCCCGCGCTCCACCCCG 1278
Db 788 CCACCAACCGCCATCATCAGAGCTGTATGAGTCCATGAGCCCGCGCTCCACCCCG 847
Qy 1279 CCAGTGTGCTGCGCCCGCCACCAATGACTCCCATCAGCATTTCTATACATCAGCGGGC 1338
Db 848 CCAGTGTGCTGCGCYACNAARUNACTCCCATCAGCATTTCTATACATCAGCGGGC 907
Qy 1339 AATAATGTGGTCTACAGCAGTACGAGGACATGCTGGTGGAGTGTGGCTGCGAGGTAG 1398
Db 908 AATAATGTGGTCTACAGCAGTACGAGGACATGCTGGTGGAGTGTGGCTGCGAGGTAG 967
Qy 1399 CGGTGCTTTCCCGCGCGCTTGGCCCG 1425
Db 968 CGGTGCTTTCCCGCGCGCTTGGCCCG 994
```

RESULT 7

AAT31602

ID AAT31602 standard; DNA; 1308 BP.

XX AAT31602;

AC AAT31602;

XX AAT31602;

25-OCT-1996 (first entry)

XX Cartilage-derived morphogenetic protein-2 gene.

DE Cattle; cartilage-derived morphogenetic protein-2; CDMP-2;

XX articular cartilage; chondrogenic; vulnery; implantation;

KW chondromalacia; osteoarthritis; therapy; joint repair; ss.

XX Bos taurus.

OS Bos taurus.

XX Location/Qualifiers

FH Key

```
FT mat_peptide 949..1308
FT /*tag= a
FT /product= Mature CDMP-2 C-terminal domain
FT 1042..1146
FT /*tag= b
FT /note= "RT-PCR-amplified probe used to screen cDNA
FT library"
XX WO9614335-A1.
PN 17-MAY-1996.
XX 17-MAY-1996.
XX 07-NOV-1994; 94WO-US12814.
XX 07-NOV-1994; 94WO-US12814.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Chang SC, Luyten FP, Moos M;
XX WPI: 1996-251714/25.
XX P-PSDB; AAR95636.
XX New purified cartilage extracts and proteins - used to stimulate the
XX development and repair of cartilage in vivo.
XX Claim 10; Fig 2; 34pp; English.
XX The sequence encodes cattle articular cartilage-derived morphogenetic
XX protein-2 (CDMP-2). Primers AAT13603-06, constructed from conserved
XX motifs from bone morphogenetic protein, have been used in reverse
XX transcription-polymerase chain reaction amplification of a cattle
XX articular chondrocyte cDNA library, generating products of 120 and
XX 280 bp, which are then used to screen a cattle articular cartilage
XX cDNA library. The 5'-end with the N-terminal methionine codon and
XX signal peptide sequence is missing. CDMP-2 is present in a purified
XX cartilage extract (claimed) which stimulates local cartilage formation
XX and repair when combined with a matrix and implanted in a mammal. The
XX protein may be used in therapy of e.g. chondromalacia or
XX osteoarthritis, to heal joint surfaces, or to repair cartilage after
XX reconstructive surgery.
```

SQ Sequence 1308 BP; 209 A; 456 C; 452 G; 191 T; 0 other;

Query Match 65.1%; Score 928.2; DB 17; Length 1308;

Best Local Similarity 86.5%; Pred. No. 3.8e-157;

Matches 1136; Conservative 0; Mismatches 123; Indels 54; Gaps 8;

```
Qy 128 COTCCGCCGAGCTGGTTCACCAAGGCGCATCGCAAGCGCGCAAGGAGCAAGATGCAGC 187
Db 5 COTCCGCCGAGCTGGTTCACCAAGGCGCATCGCAAGCGCGCAAGGAGCAAGATGCAGC 64
Qy 188 GGGCGCGCGCGGACAGTACGCGGGCGCGGGAG-----GGCCAGGACACAGCGCGC 238
Db 65 GGGCGCGCGGAGAGATGCCAGCGCGCGGGAGCCCTGGATGCGCGAGGAGCGCGCGCGCA 124
Qy 239 GGCCTCAGGACGAACCC-----CGGGCTCAGCAGCCCGCGCGCGGAGGAGCGCGCAG 289
Db 125 GGGCGCGGAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
Qy 290 GCAGGGTCCGCGCGCTGGTGGCGCCACAGTACATGCTGCAATCTACAGGACTTACTCCA 349
Db 185 GCAGGGTCCGCGCGCTGGTGGCGCCACAGTACATGCTGCAATCTACAGGACTTACTCCA 244
Qy 350 TCGCTGAGAGCTGGGCGCATCAATGCCAGCTTTTCCAGTCTTCCAGTCCGCTAATAGA 409
Db 245 TCGCGCGAGAGCTGGGCGCATCAATGCTAGCTTTTCCAGTCTTCCAGTCCGCTAATAGA 304
Qy 410 TCACGAGCTTTGTAGACAGGCGACTAGACGATCTCTCCACACTCTCTCCGAGACAGA 469
Db 305 TCACGAGCTTTGTAGACAGGCGACTAGACGATCTCTCCACACTCTCTCCGAGACAGA 364
Qy 470 AGTATTTGTTGATGTGTCATGCTCTCAGACAAAGAGAGCTGGTGGCGCGGAGCTGC 529
```



```
Db 347 GGAGCCGAGGGGTCAATGTCAGAGCGGCTCGGGCTCCCCAGACACCGGGTCTTGGCTGCC 406
QY 1012 TCGCCCGCGCGCGGGCGCGACCGCTTCGCCAGTCGCCATGCGAAGCGGCACGGC 1071
Db 407 TCGCCCGCGCGCGGGCGGACCGACCGCTTCGCCAGCGGTCCAGCGAAGCGACATGGC 466
QY 1072 AAGAAGTCCAGGCTAGCGTGCAGCAAGAAGCCCTCGACGTGAACCTCAAGGAGCTGGGC 1131
Db 467 AAGAAGTCCAGGCTAGCGTGCAGCAAGAAGCCCTCGACGTGAATTTAAGGAGTTAGGC 526
QY 1132 TGGGAGACTGGATTATCGCGGCCCTTGGAGTAGCAGAGCCCTATCACTGCGAGGGTGTATGC 1191
Db 527 TGGGAGACTGGATTATCGCGGCCCTTAGAGTAGCAGAGCCCTATCACTGCGAGGGCTGTGC 586
QY 1192 GACTTCCCGCTCGCTCGACCTGGAGCCCAACCAACAGCCCATCATCAGAGCTGATG 1251
Db 587 GACTTCCCGCTCGCTCGACCTGGAGCCCAACCAACAGCCCATCATCAGAGCGTGATG 646
QY 1252 AACTCCATGSAACCCCGGCTCCACCCCGCCAGCTGCTGGTGCCCAACCAATGACTCCC 1311
Db 647 AACTCCATGSAACCCCGGCTCCACCCCGCCAGCTGCTGGTGCCCAACCAACTGACTCCC 706
QY 1312 ATCAGCATCTATACATCAGCGGGGCAATATGTGTTACAGCAGTACAGGACATG 1371
Db 707 ATTAGCATCTGTACATCAGCGGGGCAATATGTGTTACAGCAGTACAGGACATG 766
QY 1372 GTGGTGGAGTCGCGGCTGCAGGTAGCGTCTTTCCGCGCGCTTGCCCGG 1425
Db 767 GTGGTGGAGTCGCGGCTGTAGTAGCGGTG-CTGTCCCGCCACCTGGGCCAG 819

RESULT 9
AADI8336
ID AADI8336 standard; DNA; 1046 BP.
XX
AC AADI8336;
XX
DT 18-DEC-2001 (first entry)
XX
DE Murine partial mv2 DNA.
XX
KW Mouse; bone morphogenic protein; BMP-12; vulnery; antiinflammatory;
KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;
KW tissue formation; wound healing; tissue repair; ligament defect;
KW carpal tunnel syndrome; tendonitis; mv2; ds.
XX
OS Mus sp.
XX
FH Key
FT 2..793 Location/Qualifiers
FT CDS
FT /*tag= a
FT /product= "Murine partial mv2 protein"
FT /transl_except= (pos:160..163, aa:Xaa)
FT /transl_except= (pos:740..742, aa:Xaa)
FT /note= "Xaa is an unknown amino acid; CDS does not
FT include start codon"
FT /partial
XX
PN US6284872-B1.
XX
PD 04-SEP-2001.
XX
PF 28-FEB-1997; 97US-0808324.
XX
PR 22-DEC-1994; 94US-0362670.
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
PR 02-NOV-1994; 94US-0333576.
XX
PA (GENV ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
PI Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;
```

```
PI Melton DA;
XX
DR WPI: 2001-588978/66.
DR P-PSDB; AA010985.
XX
PT New chimeric DNAs, useful for treating tendonitis, carpal tunnel
PT syndrome and other tendon and ligament defects, comprises DNA encoding
PT propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,
PT BMP-13 or MP52 -
XX
PS Example 1; Column 71-74; 42pp; English.
XX
CC The invention relates to a chimeric DNA comprising a DNA sequence
CC encoding a propeptide from a member of the transforming growth factor
CC (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence
CC encoding an amino acid sequence encoding a mature polypeptide consisting
CC of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA
CC sequences are useful for producing proteins which induce tendon/ligament
CC like tissue formation, and for isolating and cloning further DNA
CC sequences encoding BMP-12 related proteins with similar activity. The
CC proteins are useful for the induction of tendon/ligament-like tissue
CC formation, wound healing, ligament and other tissue repair, augmenting
CC the activity of bone morphogenetic proteins, and for treating tendonitis,
CC carpal tunnel syndrome and other tendon and ligament defects. The
CC present sequence is murine partial mv2 DNA which is homologous to
CC human BMP-12 or VL-1 sequences of the invention.
XX
SQ Sequence 1046 BP; 212 A; 314 C; 331 G; 187 T; 2 other;
Query Match 40.8%; Score 581.8; DB 22; Length 1046;
Best Local Similarity 87.2%; Pred. No. 2.9e-95;
Matches 675; Conservative 1; Mismatches 84; Indels 14; Gaps 3;
QY 653 CGCGCGCGCGCTGGGAAGTC-TTCGACGTGTGCGAGGGCTGGCCACAGCCCTGGAAG 711
Db 59 CCCAGCGCGCTGGGAAGTC-TTCGACGTGTGCGAGGGCTGGCCCTCAGCCCTTGGAG 118
QY 712 CAGCTGTGCTGTGAGCTGGGCGCGCATGCGGAGCTGCGAGCGGGAGCGCGG 771
Db 119 CAGCTGTGCTGTGAGTTCGGGCGAGCTGGGTGAGCTGACGCCGGGATAGCGGGCG 178
QY 772 CGCGCGCGGGGACCCAGCAACCGCGCCCGGAGCTCGGAGTCTGGGCTTCGCCCG 831
Db 179 CGCGGAGGGGTCCCGCAGCAGCACCGCCTCTGGACCTCGGAGTCTGGCTCGG 238
QY 832 AGGTGCGGCTTCCCGCAGGAGCGGCGCTGCTGGTGGTATTCCACAGATCCAGCCCAAG 891
Db 239 AGGTGAGACCGCGCGCGAGGCGCGCTGCTGTGTAGTGTTCACCAAGATCGCAGCGCAAG 298
QY 892 AACCTGTTCGACAGATGCGCGAGCTGGGCTCGGCGGAGGCTCGGCGCGCGCGCG 951
Db 299 AACCTGTTCACGTAGATGATGATGAGCAGCTGGGCTCTGACAGAGCT-----GCG 346
QY 952 GCGCGCGAGGGTCTGTGGCGCGCGCGCTCGGCGCGCGCGCGATGCCAGGCTTGGCTGCC 1011
Db 347 GGAGCGAGGGGTGATGTCCAGCGCGCTCGGCTCCCGCAGACACCGGGTCTTGGCTGCC 406
QY 1012 TCGCGCGCGCGCGGGGCGGCGAGCGGCTTCGCCAGTTCGCCATGCGGCAAGCGGCACGGC 1071
Db 407 TCGCGCGCGCGCGGCGGAGCGACCGGCTTCGCCAGCGCTTCGCCAGCGACATGGC 466
QY 1072 AAGAAGTCCAGGCTAGCTGCGAGCAAGAAGCCCTTCGACGTGAACCTCAAGGAGCTGGGC 1131
Db 467 AAGAAGTCCAGGCTAGCTGCGAGCAAGAAGCCCTTCGACGTGAATTTAAGGAGTTAGGC 526
QY 1132 TGGGAGCACTGGATTATCGCGGCCCTTGGAGTAGCAGAGCCCTATCACTGCGAGGGTGTATGC 1191
Db 527 TGGGAGCACTGGATTATCGCGGCCCTTAGATAGCAGGCGCTATCACTGCGAGGGCTGTGC 586
QY 1192 GACTTCCCGCTCGCTCGCAACCTGGAGCCCAACCAACAGCCCATCATCAGAGCTGATG 1251
Db 587 GACTTCCCGCTCGCTCGCAACCTGGAGCCCAACCAACAGCCCATCATCAGAGCTGATG 646
```

QY 1252 AACTCCATGAGACCCCGCTCCACCCCGCCAGCTGCTGCGTGCCACCAAAATTGACTCCC 1311
Db 647 AACTCCATGAGACCCCGCTCCACCCCGCCAGCTGCTGCGTGCCACCAAAATTGACTCCC 706

QY 1312 ATACAGCATTCATACATCGACGGGGGCAATAATGTGCTTACAGCAGTACGAGGACATG 1371
Db 707 ATTAGCATCTGTACATCGACGGGGGCAATAATGTGCTTACAGCAGTACGAGGACATG 766

QY 1372 GTGGTGAGTCTGCGGCTGCGAGTACGCTGCTTTCCCGCGCCCTTGCCCGG 1425
Db 767 GTGGTGAGTCTGCGGCTGCTGAGTACGGTG-CTGTCCCGCCACTTGGGCCAG 819

RESULT 10
AAQ96224
ID AAQ96224 standard; DNA; 1046 BP.
XX
AC AAQ96224;
XX
XX 23-NOV-1995 (first entry)
XX Murine mv2 protein.
XX Bone morphogenetic protein; mv2; tendon; ligament; ss.
XX Mus musculus.
XX OS
XX Location/Qualifiers
FH 2..790
FT CDS
FT /*tag= a
FT misc_difference 161
FT /*tag= b
FT /*label= C or A
FT misc_difference 742
FT /*tag= c
FT /*label= A,C,T or G
XX
XX W09516035-A.
XX
XX 15-JUN-1995.
XX
XX 06-DEC-1994; 94WO-US14030.
XX
XX 02-NOV-1994; 94US-0333576.
XX 07-DEC-1993; 93US-0164103.
XX 25-MAR-1994; 94US-0217780.
XX
XX (GENY) GENETICS INST INC.
XX (HARD) HARVARD COLLEGE.
XX
XX Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
XX Wozney JM;
XX WPI; 1995-224320/29.
XX P-PSDB; AAR78739.
XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in
XX compsn. for inducing tendon/ligament-like tissue formation
XX Example; page 71-72; 84pp; English.
XX
XX Oligos #6 and #7 (AAQ96218 & AAQ96219) are used as primers for the
XX amplification of a 275 bp DNA probe, the internal 269 bp of which
XX corresp. to nts #607 to #865 of AAQ96207, from the BMP-12 encoding
XX plasmid subclone PCR1-l#2. This probe was radioactively labelled
XX and used to screen a murine genomic library. DNA sequence analysis
XX of one of positively hybridising recombinants named MVR32 indicates
XX that it encodes a portion of the mouse gene corresp. to the PCR
XX product mv2 (murine homolog of the human VL-1 sequence AAQ96213.
XX The partial DNA sequence of this subclone and corresp. AA
XX translation are given in AAQ96224 & AAR78739.
XX
XX Sequence 1046 BP; 212 A; 314 C; 331 G; 187 T; 2 other;

Query Match 40.8%; Score 581.2; DB 16; Length 1046;
Best Local Similarity 87.2%; Pred. No. 3.7e-95;
Matches 675; Conservative 0; Mismatches 85; Indels 14; Gaps 3;
QY 653 CGCGCGCGCGTGGGAAGTC-TTGCAGCTGTGGCAGGGCCCTGCGCCACCAAGCCCTTGAAG 711
Db 59 CCCAGCGCGGTGGGAAGTCTTTCGACGTGTGGCAGGGCCCTGCGCCCTCAGCCCTTGAAG 118
QY 712 CAGCTGTGCTTGGAGCTGCGGGCCGATGGGGCAGCTGGACGCCGGGGAGCCGAGGCG 771
Db 119 CAGCTGTGCTTGGAGTTCGCGGACCTGGGGTGTGAGTGGACNCCGGGGATACGGGGGCG 178
QY 772 CGCGCGCGGGACCCAGCAACCGCCCGCCGACCTGCGGAGTCTGGGCTTGGCGCG 831
Db 179 CGCGGAGGGGTCCCCAGAGCCACCGCCCTGGACCTCGGAGTCTGGGCTTGGTCGG 238
QY 832 AGGTGCGGCTCCCCAGAGCGGGCCCTGCTGCTGTGTAATTCACAGATCCAGCGCAAG 891
Db 239 AGGTGAGACCGCCCGAGAGCGGCCCTGCTTGTAGTGTTCACAGATCGAGCGCAAG 298
QY 892 AACCTGTTCCAGAGATGCGGAGCAGCTGGGCTCGGCCGAGGCTGCGGGCCCGGCGCG 951
Db 299 AACCTGTTCACTGAGATGCAATGACGCTGGGCTCTGCAGAGGCT-----GCG 346
QY 952 GGGCGCGAGGGGTGCTGGCGCGCGCCCTCGGGCGCCCGGATGCCAGGCTTGGCTGCC 1011
Db 347 GGAGCGAGGGGTGATGTCAGAGCGCGCTCGGGCTCCCGACAGACCGGCTTGGCTGCC 406
QY 1012 TCGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCCAGTGGCCATGCAAGCGCGCAGCG 1071
Db 407 TCGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCGCCAGCGCTCACGCAAGCGACATGGC 466
QY 1072 AAGAGTCCAGGCTACGCTGCAGCAAGAGCCCTGCACGTGAATTCAGAGGAGCTGGGC 1131
Db 467 AAGAGTCCAGGCTGCGCTGCAGCAAGAGCCCTGCACGTGAATTTAAGAGGATTAGGC 526
QY 1132 TGGGACGACTGGATTATCGCGCGCCCTGGAGTACGAGGCTATCACTCGAGGGTGTATGC 1191
Db 527 TGGGACGACTGGATTATCGCGCGCCCTAGAGTACGAGGCTATCACTCGAGGGGTGTGC 586
QY 1192 GACTTCCCGCTGCGCTCGACCTGGAGCGCCACCAACAGCGCCATCATCCAGAGCGCTGATG 1251
Db 587 GACTTCCCGCTGCGCTCGACCTTGAAGCGCCACTAACCATGCCATCATTCAGACGCTGATG 646
QY 1252 AACTCCATGAGACCCCGCTCCACCCCGCCAGCTGCTGCGTGCCCGCCCAATTCAGTCCC 1311
Db 647 AACTCCATGAGACCCCGCTCCACCCCGCCAGCTGCTGCGTGCCCGCCCAATTCAGTCCC 706
QY 1312 ATCAGCATTCATACATCGACGGGGCAATAATGTGCTTACAGCAGTACGAGGACATG 1371
Db 707 ATTAGCATCTGTACATCGACGGGGCAATAATGTGCTTACAGCAGTACGAGGACATG 766
QY 1372 GTGGTGAGTCTGCGGCTGCGAGTACGGTGCGCTTTCCCGCGCCCTTGGGCCG 1425
Db 767 GTGGTGAGTCTGCGGCTGCTGAGTACGGTG-CTGTCCCGCCACTTGGGCCAG 819

RESULT 11
AAQ84280
ID AAQ84280 standard; DNA; 530 BP.
XX
AC AAQ84280;
XX
XX 17-AUG-1995 (first entry)
XX
XX GDF-6 gene.
XX
XX Growth differentiation factor-6; GDF-6;
XX transforming growth factor-beta; TGF-beta; cell proliferation;
XX pancreas; ss.
XX
XX Mus sp.


```

XX FH Key Location/Qualifiers
XX CDS 126..530
XX FT /*tag= a
XX
XX PN WO9501801-A.
XX
XX PD 19-JAN-1995.
XX
XX PF 08-JUL-1994; 94WO-US07762.
XX
XX PR 09-JUL-1993; 93US-0089300.
XX
XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MED.
XX
XX PI Huynh T, Lee S;
XX
XX DR WPI; 1995-066738/09.
XX DR P-PSDB; AAR66867.
XX
XX PT Growth differentiation factor-6 - a member of the TGF-beta
XX PT family, useful to treat cell proliferation disorders in
XX PT pancreatic tissue
XX
XX PS Disclosure; Page 31-32; 62pp; English.
XX
XX CC Primer SJL141 (given in AAQ84278), based on a region spanning 2 Trp
XX CC residues conserved in TGF-beta proteins, and primer SJL145 (AAQ84279)
XX CC spanning the invariant Cys residues near the C-terminus were used
XX CC to amplify mouse genomic DNA. New growth differentiation factor-6
XX CC (AAQ84280) was identified from a mixture of PCR products.
XX
XX SQ Sequence 530 BP; 104 A; 166 C; 160 G; 100 T; 0 other;

Query Match 28.9%; Score 412.2; DB 16; Length 530;
Best Local Similarity 87.8%; Pred. No. 5.8e-65;
Matches 477; Conservative 0; Mismatches 53; Indels 13; Gaps 2;

QY 856 GCCTGTGTGGTATTACACAGATCCAGCGCAAGAACCTGTCGAGAGATCGCGAG 915
DB 1 GGCCTGTTGTAGTGTTCACAGATCGCAGCGCAAGAACCTGTTCTAGATGCATGAG 60

QY 916 CAGCTGGGCTCGGCCAGGCTCGGGGCCCGGGCGCCGAGGGTCTGTGGCCGCCG 975
DB 61 CAGCTGGGCTCTCAGAGGCT-----GCGGGAGCGCCAGAGGGTCTATGGCCAGCG 108

QY 976 CCGTCGGGCGCCCGGATCGAGGCTTGGCTCGCCCTCGCCGCGCGCGCGCGCGCG 1035
DB 109 CCGTCGGGC-TCCAGACGCGCGGGTCTTGGCTGCGCCCTCGCCGCGCGCGCGCGCG 167

QY 1036 ACGGCCCTTCCCGAGTCGCCATGCGCAAGCGGCACGGAAGTCCAGGCTACGCTG 1095
DB 168 ACCGCCCTTCGCCAGCGCTACGCGACATGGCAAGATCCAGGCTGCGCTGCGAGC 227

QY 1096 AAGAAGCCCTGACGTGAACCTCAGGAGCTGGGCTGGGACGACATGATTCGCGGCC 1155
DB 228 AGAAGCCCTGACGTGAACCTCAGGAGCTGGGCTGGGACGACATGATTCGCGGCC 287

QY 1156 CTGGAGTACGAGCGCTATCACTCGAGGGTGTATGCGACTTCCGCTGCGCTGCGACCTG 1215
DB 288 CTAGAGTACGAGCGCTATCACTCGAGGGGCTGTGCGACTTTCCGCTGCGCTGCGACCTT 347

QY 1216 GAGCCACCAACACGCCATCATCCAGAGCGCTGATGAATCCATGACCCCGGCTCCACC 1275
DB 348 GAGCCACCAATCAATCCATTCAGAGCGCTGATGAATCCATGACCCCGGCTCCACC 407

QY 1276 CCGCCAGCTGCTGCGTCCGCCAATATGACTTCCCATCAGCATTTCTATACATGACGG 1335
DB 408 CCGCCTAGCTGCTGCTTCCGCCAATATGACTTCCCATCAGCATTTCTATACATGACGG 467

QY 1336 GGCATAATGTGCTCTACAGCAGTACGAGGACATGGTGTGAGTCTGTCGGCTCGAGG 1395
DB 468 GGCATAATGTAGTCTACAGCAGTATGAGGACATGGTGTGAGTCTGTCGGCTCGAGG 527
```

```

QY 1396 TAG 1398
DB 528 TAG 530

RESULT 12
AAA72621
ID AAA72621 standard; DNA; 530 BP.
XX
XX AAA72621;
XX
XX 29-NOV-2000 (first entry)
XX
XX Murine growth differentiation factor-6 (GDF-6) DNA sequence.
XX
XX Antibody; growth differentiation factor-6; GDF-6; placenta; mouse;
XX KW transforming growth factor-beta; TGF-beta; choriocarcinoma; neoplasm;
XX KW detection; treatment; cell proliferative disease; ds.
XX
XX OS Mus sp.
XX
XX PN US6090563-A.
XX
XX PD 18-JUL-2000.
XX
XX PF 15-JUN-1998; 98US-0097616.
XX
XX PR 15-APR-1996; 96US-0581529.
XX
XX PA (UYJO ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX
XX PI Lee S, Huynh T;
XX
XX DR WPI; 2000-531418/48.
XX DR P-PSDB; AAB12986.
XX
XX PT Antibody for detecting cell proliferative disorder associated with
XX PT growth differentiation factor-6 expression, specifically binds to
XX PT growth differentiation factor-6 polypeptide
XX
XX PS Example 2; Fig 2; 24pp; English.
XX
XX This invention relates to an antibody that specifically binds to a growth
XX differentiation factor-6 (GDF-6) polypeptide. GDF-6 is a member of the
XX transforming growth factor-beta (TGF-beta) family of proteins, and is
XX expressed in placental tissue. The antibody of the invention can be used
XX to identify tissue samples of placental origin. It may also be used to
XX treat choriocarcinomas, or detect primary and metastatic neoplasms of
XX placental origin. The antibodies may also be used as an indicator of
XX proliferative diseases, and GDF-6 may be used as an indicator of
XX developmental anomalies in prenatal screening procedures. The present
XX sequence represents DNA encoding the murine GDF-6 protein.
XX
XX SQ Sequence 530 BP; 104 A; 166 C; 160 G; 100 T; 0 other;

Query Match 28.9%; Score 412.2; DB 21; Length 530;
Best Local Similarity 87.8%; Pred. No. 5.8e-65;
Matches 477; Conservative 0; Mismatches 53; Indels 13; Gaps 2;

QY 856 GCCTGTGTGGTATTACACAGATCCAGCGCAAGAACCTGTCGAGAGATCGCGAG 915
DB 1 GGCCTGTTGTAGTGTTCACAGATCGCAGCGCAAGAACCTGTTCTAGATGCATGAG 60

QY 916 CAGCTGGGCTCGGCCAGGCTCGGGGCCCGGGCGCCGAGGGTCTGTGGCCGCCG 975
DB 61 CAGCTGGGCTCTCAGAGGCT-----GCGGGAGCGCCAGAGGGTCTATGGCCAGCG 108

QY 976 CCGTCGGGCGCCCGGATCGAGGCTTGGCTCGCCCTCGCCGCGCGCGCGCGCGCG 1035
DB 109 CCGTCGGGC-TCCAGACGCGCGGGTCTTGGCTGCGCCCTCGCCGCGCGCGCGCGCG 167

QY 1036 ACGGCCCTTCCCGAGTCCGCAAGCGGCACGGAAGTCCAGGCTACGCTGACG 1095
```



```

Db 168 ACCGCTTCGCCAGCGCTACGCGCAAGCGACATGGCAAGATCCAGGCTGCCTGCAGC 227
QY 1096 AAGAAGCCCTCGACGCTGAACCTTCAAGGAGCTGGCTGGGACGACTGGATTATCGCGGCC 1155
Db 228 AGAAGCCCTCGACGCTGAATTTTAAGGAGTTAGGCTGGGACGACTGGATTATCGCGCCC 287
QY 1156 CTGAGTACGAGCGCTATCACTACGAGGCTGTATGCGACTTCCGCTCGCCTCGCACCTG 1215
Db 288 CTAGAGTACGAGCGCTATCACTACGAGGCTGTATGCGACTTCCGCTCGCCTCGCACCTT 347
QY 1216 GAGCCCAACCAACGCGCTATCACTACGAGGCTGTATGCGACTTCCGCTCGCCTCGCAC 1275
Db 348 GAGCCCACTAACCATCCATTCATCAAGCGCTGATGAACCTCCATGAGCCCGGCTCCACC 407
QY 1276 CCGCCAGCTGCTGCGTGGCCCAAAATGACTCCCATCAGCATTTCTATACATCGAGCG 1335
Db 408 CCGCCAGCTGCTGCGTGGCCCAAAATGACTCCCATCAGCATTTCTATACATCGAGCG 467
QY 1336 GGCAATAATGTGCTTACAAAGCAGTACGAGGACATGGTGGTGGAGTCGTGCGCTGCAGG 1395
Db 468 GGCAATAATGTGCTTACAAAGCAGTATGAGGACATGGTGGTGGAGTCCTGCGCTGTAGG 527
QY 1396 TAG 1398
Db 528 TAG 530

RESULT 13
AAQ96215
ID AAQ96215 standard; DNA; 1345 BP.
XX
AC AAQ96215;
XX
22-NOV-1995 (first entry)
XX
DE Human bone morphogenetic protein (BMP)-12 DNA.
XX
XX Bone morphogenetic protein; BMP-12; tendon; ligament; tendinitis; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 138..1301
FT CDS /*tag= a
FT mat_peptide 990..1301
FT /*tag= b
XX
PN W09516035-A.
XX
15-JUN-1995.
XX
06-DEC-1994; 94WO-US14030.
XX
02-NOV-1994; 94US-0333576.
PR 07-DEC-1993; 93US-0164103.
PR 25-MAR-1994; 94US-0217780.
XX
(GEMY ) GENETICS INST INC.
PA (HARD ) HARVARD COLLEGE.
XX
Celeste AJ, Melton DA, Rosen VA, Thomsen GH, Wolfman NM;
PI Wozney JM;
XX
XX WPI; 1995-224320/29.
DR P-PSDB; AAR78734.
XX
XX Bone morphogenetic proteins -12 and -13 and corresp. DNA - used in
PT compsn. for inducing tendon/ligament-like tissue formation
XX
XX Example; Page 73-75; 84pp; English.
XX
XX Oligo probe 3 (AAQ96214) designed on the basis of an amplified BMP-12
CC
```

```

CC human DNA sequence (AAQ96212) was radioactively labelled and used to
CC screen a human genomic library. A positive clone was designated
CC lambda-HuG-48. This bacteriophage was deposited under ATCC 75625 on 7
CC Dec. 1993. The oligo hybridising region of this phage is localised
CC to a 3.2 kb BamHI fragment which was subcloned to PCR1-1#2 and
CC deposited under ATCC 69517 on 7 Dec. 1993. The partial DNA sequence
CC and derived AA sequence of the 3.2 kb DNA insert of the plasmid
CC subclone PCR1-1#2, derived from clone lamdba-HuG-48 are shown in
CC AAQ96207 and AAR78729 respectively. Additional DNA sequence of the
CC 3.2kb BamHI insert of the plasmid PCR1-1#2 is given in AAQ96215.
CC All of AAQ96207 is contained in AAQ96215. As AAQ96215 is derived from
CC a genomic clone it is difficult to determine the boundary between
CC the 5' extent of coding sequence and the 3' limit of intervening
CC sequence.
XX
SQ Sequence 1345 BP; 205 A; 447 C; 475 G; 218 T; 0 other;
Query Match 18.2%; Score 259.2; DB 16; Length 1345;
Best Local Similarity 57.8%; Pred. No. 1.2e-37;
Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;
QY 464 GACAGAAGTATTTTGTGATGTCTCCATGCTCTCAGACAAAGAAGACCTGGTGGCGCGG 523
Db 364 GCCAGAGCTTCCCTGTTCGACGTGTCCAGCCTTAACGACGACAGAGGTGGTGGTCCG 423
QY 524 AGCTGGCGCTCTTTTCGCCAGCGCCCTCAGCGCCCTGGGGGCCACACAGCGCGCGCTCC 583
Db 424 AGCTGGCGCTGTGCGCGGGATCTCCAGAGTC--GGGCCAGGACAGCTGGACTTCTCC 481
QY 584 ACCTGACGCTCTTCCCTTTCGCTTTCGCTTCTACTGTGACGCGGAGACCTTGACACCGC 643
Db 482 -----GCCGTTGCTGCTGTCCACGTGCGCGGCGCGCGCGACGCCACGCGCTGC 534
QY 644 AGGGGGCGCGCGCGCGCTGGGAAGTCTTCGACGTGTGGGAGGCGCTGGCCACCACG 703
Db 535 TGTACTCGCGGGGACGTGAGCCCTAGTCGGTCAAGCGCTGGGAGCGGTTTCAGCGTGCGCG 594
QY 704 CTGGAAGACAGCTGTGCTTGGAGCTGCGGGCCGATGGGCGGAGCTGGAGCGCGGGGAGG 763
Db 595 ACGCCATGAGCGCCACCGCTCGTAACCGCGCCCGCCCGCGCTTCTGCTCTTCTGCTGC 654
QY 764 CCGAGCGCGCGCGCGGGACCCAGCAACCGCGCGCGCGCGCGCGCGGAGTGTGGGCT 823
Db 655 GCGCAGTGGCAGCGCGCGGTGCCAGCCGTTGGCACTTGGCGGACCTGGGCTTGGGCTGC 714
QY 824 TCGGCGGAGGCTGCGCGCTCCCGCAGGAGCGCGCGCTGTGTTGGTATTACCATGATCCC 883
Db 715 CGGCGGAGGGGCTCTGCGGCAGAGAGCGCGGCTAGTCTCTCTCCCGCACGC 774
QY 884 AGCGCA---AGAACCTGTTTCGACAGATGCGGACGAGCTGGGCTCGCGGAGGCTGC-- 938
Db 775 AGAGGAAGAGAGCTTATTCGGGAGATCGCGCCAGGCGCGCGCGCGCGCGCTC 834
QY 939 -----GGGCCCGGGCGCGGGCGCGAGGGGTTGTTGGCGCGCGCGCGCGCGCGCC 988
Db 835 TGGCCTCAGAGCGCTGCGCGACCCAGGAACCGGCACCGCTGCGCAAGGCGAGTCATTG 894
QY 989 CGGATGCCAGGCTTGGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGC 1048
Db 895 GCGGCCGACAGAGGAGGAGCGCGCTTGGCGGGGACCGCGGACAGCGGAGCGCGCGG 954
QY 1049 GTCGCCATGGCAAGCGGCACGCAAGAAGTCCAGGCTACGCTGCAGCAAGAAGCCCTGC 1108
Db 955 GGGCGCGGGCGCGGGCCACCGCGCGAGGGCGCGGAGCCCTGCAGCGCAAGCCGTTC 1014
QY 1109 ACCTGAACCTTCAAGGAGCTGGGCTGGGACGACTGGATTATTCGCGCGCGCTGGAGTACGAG 1168
Db 1015 ACCTGGACTTCAAGGAGCTGGGCTGGGACGACTGGATCATCGCGCGCTGGAGTACGAG 1074
QY 1169 CTTATCACTGCGAGGCTGTATGCGACTTCCCGCTGCGCTGCGACCTGGAGGCCCAAC 1228
Db 1075 CGTACCACTGCGAGGCGCTTTGCGACTTTCCTTTGGTTCGACCTCGAGCCCCACCAACC 1134
```


Db 1255 TCTACAAGCAATACGAGCATGGTGTGGAGCGCTGCGCTGCAGGTAGCG 1306

RESULT 15

AD18337
ID AAD18337 standard; DNA; 1345 BP.

XX AC

XX AC

XX DT

XX DT

XX 18-DEC-2001 (first entry)

XX DE

XX Human bone morphogenic protein-12 (BMP-12) DNA #2.

XX KW

XX Human; bone morphogenic protein; BMP-12; vulnary; antiinflammatory;

XX KW analgesic; gene therapy; transforming growth factor-beta; TGF-beta;

XX KW tissue formation; wound healing; tissue repair; ligament defect;

XX KW carpal tunnel syndrome; tendonitis; ds.

XX OS

XX Homo sapiens.

XX FH

XX Key

XX CDS

XX Location/Qualifiers

XX 138..1304

XX /tag= a

XX /product= "Human BMP-12 protein"

XX /note= "CDS does not include start codon"

XX /partial

XX sig_peptide

XX 138..989

XX /tag= b

XX mat_peptide

XX 990..1301

XX /tag= c

XX /product= "Mature BMP-12 protein"

XX US6284872-B1.

XX 04-SEP-2001.

XX 28-FEB-1997; 97US-0808324.

XX 22-DEC-1994; 94US-0362670.

XX 07-DEC-1993; 93US-0164103.

XX 25-MAR-1994; 94US-0217780.

XX 02-NOV-1994; 94US-0333576.

XX (GENY) GENETICS INST INC.

XX (HARD) HARVARD COLLEGE.

XX Celeste AJ, Wozney JM, Rosen VA, Wolfman NM, Thomsen GH;

XX PI Melton DA;

XX WPI; 2001-588978/66.

XX P-PSDB; AAE10986.

XX New chimeric DNAs, useful for treating tendonitis, carpal tunnel

XX syndrome and other tendon and ligament defects, comprises DNA encoding

XX propeptide linked to DNA encoding bone morphogenetic proteins (BMP)-12,

XX BMP-13 or MP52 -

XX Example 1; Column 73-78; 42pp; English.

XX The invention relates to a chimeric DNA comprising a DNA sequence

XX encoding a propeptide from a member of the transforming growth factor

XX (TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence

XX encoding an amino acid sequence encoding a mature polypeptide consisting

XX of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA

XX sequences are useful for producing proteins which induce tendon/ligament

XX like tissue formation, and for isolating and cloning further DNA

XX sequences encoding BMP-12 related proteins with similar activity. The

XX proteins are useful for the induction of tendon/ligament-like tissue

XX formation, wound healing, ligament and other tissue repair, augmenting

XX the activity of bone morphogenetic proteins, and for treating tendonitis,

XX carpal tunnel syndrome and other tendon and ligament defects. The

XX present sequence is a DNA encoding human BMP-12 protein.

XX SQ Sequence 1345 BP; 205 A; 447 C; 475 G; 218 T; 0 other;

Query Match 18.2%; Score 259.2; DB 22; Length 1345;

Best Local Similarity 57.8%; Pred. No. 1.2e-37;

Matches 550; Conservative 0; Mismatches 378; Indels 24; Gaps 4;

QY 464 GACAGAAGTATTGTTGATGTGTCATGCTCTACACAAGAAGAGCTGGTGGCGCGG 523

Db 364 GCCAGAGCTTCTGTCGACGTGTCACGCTTAACACGACGACGAGGTGGTGGTGGCG 423

QY 524 AGTGGGCTCTTTTCGCCAGGCGCCTCAGCGCCTTGGGGGCCACACAGCGGCGCTCC 583

Db 424 AGTGGGCTCTTTTCGCCAGGCGCCTCAGCGCCTTGGGGGCCACACAGCGGCGCTCC 481

QY 584 AGTGGGCTCTTTTCGCCAGGCGCCTCAGCGCCTTGGGGGCCACACAGCGGCGCTCC 643

Db 482 -----GCCGTTGCTGCTGCTCCACGTGCGCGCGCGCGCGCGCGCGCGCGCGCTGC 534

QY 644 AGGGGGCGCGCGCGCGCGCGCTGCGAAGTCTTCACGCTGTGGCAGGCGCTTCCGCCACACG 703

Db 535 TGTACTCGCGCGCAGCTGAGCCCTTAGTCTGGTTCAGCGCTGGGAGGCGTTTCGAGTGGCGG 594

QY 704 CTTGGAAGCAGCTGCTTTGGAGCTGCGGGCCGCTATGGGCGGAGCTGGACGCGGGGAGG 763

Db 595 ACGCCATGAGGCGCCACCGCTGCTGAACCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGC 654

QY 764 CCGAGCT 823

Db 655 GCGCAGTGGCAGCTG 714

QY 824 TCGGCGGAGGCTGCGCGCTCCCGAGGAGCGGCGCGCTGCTGGTGTGTTATCACCAGATCC 883

Db 715 CGGGCGGAGGCGCGCTGCGGCGGAGGAGCGCGCGGTGCTAGTCGCTCTCTCCCGACGC 774

QY 884 AGCGCA---AGAACCTGTTTCGACAGATGCGGAGCAGCTGGGCTCGGCCGAGGCTGC-- 938

Db 775 AGAGGAAAGAGAGCTTATTTCGGGAGATCCGCGCCAGCGCGCGCGCGCTCGGGGCGCGTC 834

QY 939 -----GGGCG 988

Db 835 TGGCGCTCAGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 894

QY 989 CGGATGCCAGGCTTGGCTGCCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1048

Db 895 GCGGCGCGCAGCGGAGGAGCGCGCTTGGCGGAGCGCGGAGCGGAGCGGAGCGGAGCGGCG 954

QY 1049 GTGCGCATGGCAAGCGGCTG 1108

Db 955 GGGGCG 1014

QY 1109 AGGTGAACCTTCAAGGAGCTGGGCTGGGAGCTGGGATTCGCGGCGCGCGCGCGCGCGCG 1168

Db 1015 AGGTGAACCTTCAAGGAGCTGGGCTGGGAGCTGGGATTCGCGGCGCGCGCGCGCGCGCG 1074

QY 1169 CCTATCAGTGGGAGGCTGTATGCGACTTCCCGTGGGCTGCGACCTTGGAGCGCGCGCGCG 1228

Db 1075 CGTACACCTGCGGAGGCGCTTGGGACTTCCCTTGGCTTGGCAGCTTGGAGCGCGCGCGCG 1134

QY 1229 AGCCCATCATCAGAGCGCTGATGAATCCATCCATCCCGCGCGCGCGCGCGCGCGCGCG 1288

Db 1135 ATGCCATCATCAGAGCGCTGCTCACTCCATCCATCCCGCGCGCGCGCGCGCGCGCGCG 1194

QY 1289 GGTGCG 1348

Db 1195 GTGTGCG 1254

QY 1349 TCTACAAGCAGTACGAGGAGCATGGTGGTGGAGTGGTGGGCTGCGGCGCTGCGAGGTAGCG 1400

Db 1255 TCTACAAGCAGTACGAGGAGCATGGTGGTGGAGGCGCTTGGCGCTGCGAGGTAGCG 1306

Job time : 324 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:52:20 ; Search time 51 Seconds
(without alignments)
857.670 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLISFLMDL.....GNNVYKQYEDMVVESCGR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1798	73.5	436	2 B55452	cartilage-derived
2	927.5	37.9	501	2 JC2347	growth/differentia
3	919.5	37.6	495	2 S43294	bone morphogenetic
4	918.5	37.5	501	2 A55452	cartilage-derived
5	691	28.2	125	2 S43295	bone morphogenetic
6	549	22.4	151	2 S43296	bone morphogenetic
7	470	19.2	398	2 JH0687	bone morphogenetic
8	464	19.0	398	2 JH0688	bone morphogenetic
9	453	18.5	405	2 I50608	bone morphogenetic
10	441.5	18.0	408	1 BMH4	bone morphogenetic
11	436.5	17.8	420	2 I49541	bone morphogenetic
12	434	17.7	393	2 S37073	bone morphogenetic
13	432.5	17.7	408	2 S38343	bone morphogenetic
14	431.5	17.6	396	1 BMH2	bone morphogenetic
15	429.5	17.6	461	2 S52408	SPDVR1 protein - s
16	426.5	17.4	408	2 JH0801	bone morphogenetic
17	423	17.3	394	2 S45355	bone morphogenetic
18	419.5	17.1	408	2 S58791	bone morphogenetic
19	418.5	17.1	452	2 I49542	bone morphogenetic
20	417.5	17.1	353	2 I50607	bone morphogenetic
21	416.5	17.0	402	2 A45056	osteogenic protein
22	416.5	17.0	454	1 BMH5	bone morphogenetic
23	413.5	16.9	401	2 JH0689	bone morphogenetic
24	408.5	16.7	400	2 A49147	bone morphogenetic
25	407	16.6	426	2 JH0590	bone morphogenetic
26	401.5	16.4	431	1 BMH7	bone morphogenetic
27	401	16.4	430	2 JQ1184	osteogenic protein
28	395.5	16.2	588	2 A26158	decapentaplegic pr
29	394.5	16.1	510	2 A54798	Vg-1-related prote

30	394.5	16.1	513	1 BMH6	bone morphogenetic
31	391	16.0	313	2 IS1284	bone morphogenetic
32	377	15.4	354	2 S29718	gene nodal protein
33	372	15.2	427	2 A40735	TGF beta homolog d
34	351	14.3	476	2 JC4646	bone morphogenetic
35	350.5	14.3	478	2 JC4838	bone morphogenetic
36	343	14.0	365	2 T43286	cet-1 protein - Ca
37	343	14.0	455	2 A43918	TGF-beta-related p
38	334	13.6	207	2 S37618	vgr protein - rat
39	325.5	13.3	472	1 BMH3	bone morphogenetic
40	324.5	13.3	360	2 A29619	Vgl embryonic grow
41	319.5	13.1	372	2 C39364	GDF-1 embryonic gr
42	318	13.0	426	1 B24248	inhibin beta-A cha
43	313.5	12.8	425	2 I47072	inhibin beta-A cha
44	310	12.7	424	1 WFGBA	inhibin beta-A cha
45	306.5	12.5	425	1 S50898	inhibin beta-A cha

ALIGNMENTS

RESULT 1

B55452

cartilage-derived morphogenetic protein 2 precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 26-May-2000
C:Accession: B55452

R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko
J. Biol. Chem. 269, 28227-28234, 1994
A:Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr
A:Reference number: A55452; MUID:95050604; PMID:7961761
A:Accession: B55452
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-436 <CHA>
A:Cross-references: GB:U13661; NID:9632489; PIDN:AAA61416.1; PID:9632490
C:Superfamily: inhibin

Query Match 73.5%; Score 1798; DB 2; Length 436;

Best Local Similarity 80.8%; Pred. No. 8.1e-120;

Matches 354; Conservative 18; Mismatches 48; Indels 18; Gaps 8;

QY	33	SSAELGSTRKGRSRKEGKQMRAPRDSADGRE---	GGEPPQPRQDEPR---	AQOPRAQEP	86
DB	2	ASAEELGSGKMRTRKEGRMPRAPRENATAREPLDR	QEPPEPPQEPPEPPQEPPEPPQEPPEPPQEP	PEARPP	61
QY	87	GRGPVVPVPEHYMLSIYRTYSIAEKLGINASFFOSS	SANTITTSFVDRGLDLSHTPLRRQ		146
DB	62	GRGPRLVPHEYMLSIYRTYSIAEKLGINASFFOSS	SANTITTSFVDRGLDLSHTPLRRQ		121
QY	147	KYLFVYSMLSDKEELVGAELRLFRQAPSAPWPPAG	PLHVQLFPCLSPDLLDARTLDPOG		206
DB	122	KYLFVYSMLSDKEELVGAELRVLRQAPALAPPAA	APLAALRLP-VAPAGSREP-GPAG		179
QY	207	APPAGVEFVDVWQGLRHQPMKQCLCLLRAAW-GE	LDAAGEAEARAGPQQPPPPDLRLSGF		265
DB	180	APRPGWEFVDVWRGLRQPWKQCLCLLRAAWGCE	PGAEAEARTPGPQQPPPPDLRLSGF		239
QY	266	GRVRPQERALLVVTTRSQKNLFAEMREQLGSA-EA	AGPGAGAGSGW-----PPPS		317
DB	240	GRVVRTPOERALLVVFVRSQRKTLFAEMREQLGS	ATVFWPGGGAGSGSGPPPPPPPPPS		299
QY	318	GAPDARFPLVSPGRRRRRTAFASRHGKSRKSLR	CKSKPLHVNFKELGWDWIAPLE		377
DB	300	GTDPAGLWSPG-RRRTAFASRHGKSRKSLRCKSK	PLHVNFKELGWDWIAPLE		358
QY	378	YEAYHCEGVCDPFLRSHLEPTNHAIITQLMNSMD	PGSTPPSCCVPTKLTPIISILYIDAGN		437
DB	359	YEAYHCEGVCDPFLRSHLEPTNHAIITQLMNSMD	PGSTPPSCCVPTKLTPIISILYIDAGN		418
QY	438	NVYKQYEDMVVESCGR	455		
DB	419	NVYINEYEMVWESCGR	436		

RESULT 2

JC2347
 growth/differentiation factor 5 - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Mar-2000
 C:Accession: JC2347
 R:Hoetten, G.; Neidhardt, H.; Jacobowsky, B.; Pohl, J.
 Biochem. Biophys. Res. Commun. 204, 646-652, 1994
 A:Title: Cloning and expression of recombinant human growth/differentiation factor 5.
 A:Reference number: JC2347; MUID:95071375; PMID:7980526

A:Accession: JC2347
 A:Molecule type: DNA
 A:Residues: 1-501 <HOE>
 A:Cross-references: GB:X80915; NID:9671524; PIDN:CAA56874.1; PID:9671525
 C:Genetics:
 A:Gene: GDB:BMP9
 A:Cross-references: GDB:433948
 A:Transons: 211/1

Keywords: glycoprotein
 Superfamily: inhibin

F:189/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:381-382/Cleavage site: Arg-Ala (unidentified proteinase) #status predicted

Query Match 37.9%; Score 927.5; DB 2; Length 501;
 Best Local Similarity 44.7%; Pred. No. 3.6e-58;
 Matches 209; Conservative 64; Mismatches 96; Indels 97; Gaps 16;

QY 38 GSTKGMRSRKEGKQAPRDSADAGREGQPPQPRQDEPRAQO----- 79
 DB 81 GOTGGLTQPKDEPKLP-----PRPG-GPEPKPGHPQTPQRTQARTVTPKGQPGGRAP 134
 QY 80 -----PRAQEPGPGPRVPHYHEMLSYRTYSTAEKLGINASFFQ 119
 DB 135 PRAGSVSPSFLKKAKEPGPPREPKEPFPPTTTPHEYMLSYRTLSADRGKGNSSVKL 194
 QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKDELVLFRQAPS----- 174
 DB 195 EAGLANTITFDKGGDDRGV-VKQKRVYFDISAL-EKDGGLGELRLRKPSDTPAKP 252
 QY 175 -APWGPAGPLHVQLPCLLS-----PLLLDARTLDPGAPAGWEVDMQGLRH-QPMWQ 228
 DB 253 AAPGGGAAQL--KLSSCPGSRQAPALLDVRVS--PGLDGSQWEVFDIWKLFERNKNSAQ 308
 QY 229 LCLELRAANGELDAGEAEARAGPQPPPDRLSLGFGRRVPPQERALLVFTSRQKN 288
 DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRAARQVHEKALFLVFGRTKKRD 352
 DB 289 L-FAEMREOLGSAEAGPCAGAGSWPPSPGAPDARPWLPSPGRRRTAFASRHKRGH 347
 DB 353 LFNELKARSGQDDKT-----VVEYLFQSO-RRKRRAPLATROGKRPS 393
 QY 348 KRSKLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSHLEPTNHAIQTILM 407
 DB 394 KNLKARCSKRALHVNFKDMGWDDWIIAPLEYEAFHCEGLCEPFLRSHLEPTNHAIQTILM 453
 QY 408 NSMDGSPTPSCCVTKLTPISILIIDAGNNVVYKQYEDMVMVESGCCR 455
 DB 454 NSMDPESTPTCCVTRLSPISILFIDSANNVVYKQYEDMVMVESGCCR 501

RESULT 3

S43294
 bone morphogenetic protein-related protein (GDF5) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 17-Mar-2000
 C:Accession: S43294
 R:Storm, E.F.; Huyh, T.V.; Copeland, N.G.; Jenkins, N.A.; Kingsley, D.M.; Lee, S.J.
 Nature 368, 639-643, 1994
 A:Title: Limb alterations in brachypodism mice due to mutations in a new member of the
 A:Reference number: S43294; MUID:94195427; PMID:8145850
 A:Accession: S43294

A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-495 <STO>
 A:Cross-references: GB:U08337; NID:9488461; PIDN:AAA18778.1; PID:9488462
 C:Superfamily: inhibin

Query Match 37.6%; Score 919.5; DB 2; Length 495;
 Best Local Similarity 43.5%; Pred. No. 1.3e-57;
 Matches 209; Conservative 74; Mismatches 107; Indels 91; Gaps 18;

QY 21 PGFOQASISSCSAELGSGTGMRSRKEGKQAPRDSADAGREGQPPQPRQDE----- 74
 DB 60 PGHHYGVGATNARAK-GSSGQTQAKDEPRKMPRSG-----GSETKPGPSSQTRQAAA 113
 QY 75 ----PRAQEP-----RAQEP-----PGRGPRVPHYHEMLSYRT 104
 DB 114 RTVTPKGQLPGKASSKAGSAPSSFLKKTRPPGTPREPKEPFPPTTTPHEYMLSYRT 173
 QY 105 YSIAEKLGINASFFOSKSNANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKDELVLGA 164
 DB 174 LSDADRKGNSSVKLEAGLANTITSFIDKGGDDRG-PAVRKQRYVFDISAL-EKDGLLGA 231
 QY 165 ELRLFRQAP---SAPWGPAGPL-HVQLPCLLS-----PLLLDARTLDPGAPAGWEVFD 216
 DB 232 ELRLRKPLDVAKPAPVSSGRVAQLKLSSCPGSRQAPALLDVRVS--PGLDGSQWEVFD 289
 QY 217 VWQGLRH-QPMWQGLRH-QPMWQGLRH-QPMWQGLRH-QPMWQGLRH-QPMWQGLRH 275
 DB 250 TWKLFERNKNSAQLCLELE-AW---ERGRA-----VDLRGLGFDERTARQVHEK 333
 QY 276 ALLVYFTSRQKNL-FAEMREOLGSAEAGPCAGAGSWPPSPGAPDARPWLPSPGRRRR 334
 DB 334 ALFLVFGRTKKRDLEFNELKARSGQDDKT-----VVEYLFQSO-RRKR 374
 QY 335 RTAFASRHKRGHKKSRKLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSH 394
 DB 375 RAPLANRQGRKPSKNLKARCSKRALHVNFKDMGWDDWIIAPLEYEAFHCEGLCEPFLRSH 434
 QY 395 LEPTNHAIQTILMSMDGSPTPSCCVTKLTPISILIIDAGNNVVYKQYEDMVMVESGCCR 454
 DB 435 LEPTNHAIQTILMSMDPESTPTCCVTRLSPISILFIDSANNVVYKQYEDMVMVESGCCR 494
 QY 455 R 455
 DB 495 R 495

RESULT 4

A55452
 cartilage-derived morphogenetic protein 1 precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 17-Mar-2000
 C:Accession: A55452
 R:Chang, S.C.; Hoang, B.; Thomas, J.T.; Vukicevic, S.; Luyten, F.P.; Ryba, N.J.P.; Ko
 J. Biol. Chem. 269, 28227-28234, 1994
 A:Title: Cartilage-derived morphogenetic proteins. New members of the transforming gr
 A:Reference number: A55452; MUID:95050604; PMID:7961761

A:Accession: A55452
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-501 <CHA>
 A:Cross-references: GB:U13660; NID:9600731; PID:9600732
 C:Genetics:
 A:Gene: GDB:CDMP1
 A:Cross-references: GDB:438940
 C:Superfamily: inhibin

Query Match 37.5%; Score 918.5; DB 2; Length 501;

Best Local Similarity 44.3%; Pred. No. 1.5e-57;
 Matches 207; Conservative 64; Mismatches 101; Indels 95; Gaps 15;

QY 38 GSTKGMRSRKEGKQAPRDSADAGREGQPPQPRQDEPRAQO----- 79
 DB 81 GOTGGLTQPKDEPKLP-----PRPG-GPEPKPGHPQTPQRTQARTVTPKGQPGGRAP 134
 QY 80 -----PRAQEPGPGPRVPHYHEMLSYRTYSTAEKLGINASFFQ 119
 DB 135 PRAGSVSPSFLKKAKEPGPPREPKEPFPPTTTPHEYMLSYRTLSADRGKGNSSVKL 194
 QY 120 SSKSANTITSFVDRGLDLSHTPLRRQKYLFDVSMLSKDELVLFRQAPS----- 174
 DB 195 EAGLANTITFDKGGDDRGV-VKQKRVYFDISAL-EKDGGLGELRLRKPSDTPAKP 252
 QY 175 -APWGPAGPLHVQLPCLLS-----PLLLDARTLDPGAPAGWEVDMQGLRH-QPMWQ 228
 DB 253 AAPGGGAAQL--KLSSCPGSRQAPALLDVRVS--PGLDGSQWEVFDIWKLFERNKNSAQ 308
 QY 229 LCLELRAANGELDAGEAEARAGPQPPPDRLSLGFGRRVPPQERALLVFTSRQKN 288
 DB 309 LCLELE-AW---ERGRA-----VDLRGLGFDRAARQVHEKALFLVFGRTKKRD 352
 DB 289 L-FAEMREOLGSAEAGPCAGAGSWPPSPGAPDARPWLPSPGRRRTAFASRHKRGH 347
 DB 353 LFNELKARSGQDDKT-----VVEYLFQSO-RRKRRAPLATROGKRPS 393
 QY 348 KRSKLRCSKKPLHVNFKELGWDWIIAPLEYEAYHCEGVCDPFLRSHLEPTNHAIQTILM 407
 DB 394 KNLKARCSKRALHVNFKDMGWDDWIIAPLEYEAFHCEGLCEPFLRSHLEPTNHAIQTILM 453
 QY 408 NSMDGSPTPSCCVTKLTPISILIIDAGNNVVYKQYEDMVMVESGCCR 455
 DB 454 NSMDPESTPTCCVTRLSPISILFIDSANNVVYKQYEDMVMVESGCCR 501

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 02:54:00 ; Search time 36 Seconds
(without alignments)
371.873 Million cell updates/sec

Title: US-09-825-751A-20

Perfect score: 2447

Sequence: 1 MDTPRVLLSAVFLSLFMDL.....GNNVYKQYEDMVESGCR 455

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1747	71.4	321	1	US-08-362-670B-26
2	1747	71.4	321	3	US-08-333-576C-26
3	1747	71.4	321	4	US-08-808-324-26
4	1747	71.4	321	5	PCT-US94-14030A-26
5	1202.5	49.1	263	1	US-08-362-670B-32
6	1202.5	49.1	263	3	US-08-333-576C-32
7	1202.5	49.1	263	4	US-08-808-324-32
8	1202.5	49.1	263	5	PCT-US94-14030A-32
9	927.5	37.9	501	2	US-08-288-508C-2
10	919.5	37.6	495	1	US-08-455-559-10
11	919.5	37.6	495	4	US-09-145-060-10
12	919.5	37.6	495	5	PCT-US94-00657-10
13	917.5	37.5	401	3	US-08-289-222E-3
14	917.5	37.5	401	4	US-09-054-526B-3
15	765	31.3	388	1	US-08-362-670B-34
16	765	31.3	388	3	US-08-333-576C-34
17	765	31.3	388	4	US-08-808-324-34
18	765	31.3	388	5	PCT-US94-14030A-34
19	731	29.9	134	1	US-08-581-529B-6
20	731	29.9	134	3	US-09-097-616-6
21	731	29.9	134	5	PCT-US94-07762-6
22	691.5	28.3	294	1	US-08-362-670B-2
23	691.5	28.3	294	3	US-08-333-576C-2
24	691.5	28.3	294	4	US-08-808-324-2
25	691.5	28.3	294	5	PCT-US94-14030A-2
26	661	27.0	119	1	US-08-581-529B-7
27	661	27.0	119	3	US-09-097-616-7

28	661	27.0	119	5	PCT-US94-07762-7	Sequence 7, Appli
29	627	25.6	411	1	US-08-362-670B-28	Sequence 28, Appl
30	627	25.6	411	3	US-08-333-576C-28	Sequence 28, Appl
31	627	25.6	411	4	US-08-808-324-28	Sequence 28, Appl
32	627	25.6	411	5	PCT-US94-14030A-28	Sequence 28, Appl
33	618	25.3	240	1	US-08-362-670B-30	Sequence 30, Appl
34	618	25.3	240	3	US-08-333-576C-30	Sequence 30, Appl
35	618	25.3	240	4	US-08-808-324-30	Sequence 30, Appl
36	618	25.3	240	5	PCT-US94-14030A-30	Sequence 30, Appl
37	553	22.6	161	2	US-08-581-528A-6	Sequence 6, Appli
38	553	22.6	161	5	PCT-US94-07799-6	Sequence 6, Appli
39	549	22.4	119	1	US-08-455-559-13	Sequence 13, Appl
40	549	22.4	119	4	US-09-145-060-13	Sequence 13, Appl
41	549	22.4	119	5	PCT-US94-00657-13	Sequence 13, Appl
42	549	22.4	120	1	US-08-362-670B-4	Sequence 4, Appli
43	549	22.4	120	3	US-08-333-576C-4	Sequence 4, Appli
44	549	22.4	120	4	US-08-808-324-4	Sequence 4, Appli
45	549	22.4	120	5	PCT-US94-14030A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-362-670B-26
; Sequence 26, Application US/08362670B
; Patent No. 5658882
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,670B
; FILING DATE: December 22, 1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-670B-26

Query Match 71.4%; Score 1747; DB 1; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRQRYKLYFDVSMLSKDELVGAEIRLFRQAPSAPWGPPAGPLHVQLFPCLSPILL 196
|||||

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAGPLHVQLFPCLSPLL 62
QY 197 LDARTLDPOGAPPAGWEVFDVWQGLRHQPKWOLCLLELRAAWGELDAGEAEARAGPQQPP 256
Db 63 LDARTLDPOGAPPAGWEVFDVWQGLRHQPKWOLCLLELRAAWGELDAGEAEARAGPQQPP 122
QY 257 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 316
Db 123 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 182
QY 317 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKKPLHVNFKELGWDWIIAPL 376
Db 183 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKKPLHVNFKELGWDWIIAPL 242
QY 377 EYAYHCEGVCDFFLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 436
Db 243 EYAYHCEGVCDFFLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 302
QY 437 NNVYKQYEDMVVESCGR 455
303 NNVYKQYEDMVVESCGR 321

RESULT 2

US-08-333-576C-26
; Sequence 26, Application US/08333576C
; Patent No. 6027919
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,576C
; FILING DATE: No. 6027919ember 2, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-576C-26

Query Match 71.4%; Score 1747; DB 3; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAGPLHVQLFPCLSPLL 196
|||||

Db 3 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAGPLHVQLFPCLSPLL 62
QY 197 LDARTLDPOGAPPAGWEVFDVWQGLRHQPKWOLCLLELRAAWGELDAGEAEARAGPQQPP 256
Db 63 LDARTLDPOGAPPAGWEVFDVWQGLRHQPKWOLCLLELRAAWGELDAGEAEARAGPQQPP 122
QY 257 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 316
Db 123 PDLRLSLGFGRRVRPPQERALLVVFTRSORKNLFAEMREQLGSAAAGPGAGAGSGWPPP 182
QY 317 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKKPLHVNFKELGWDWIIAPL 376
Db 183 SGAPDARWLPSPGRRRRRTAFASRHGKRGKSKRLRCSKKPLHVNFKELGWDWIIAPL 242
QY 377 EYAYHCEGVCDFFLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 436
Db 243 EYAYHCEGVCDFFLRSHLEPTNHAIQTLMNSMDPGSTPPSCCVPTKLTPIISILYIDAG 302
QY 437 NNVYKQYEDMVVESCGR 455
303 NNVYKQYEDMVVESCGR 321

RESULT 3

US-08-808-324-26
; Sequence 26, Application US/08808324
; Patent No. 6284872
; GENERAL INFORMATION:
; APPLICANT: Celeste, Anthony J.
; APPLICANT: Wozney, John
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wolfman, Neil
; APPLICANT: Thomsen, Gerald H.
; APPLICANT: Melton, Douglas A.
; TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENETICS INSTITUTE, INC.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,324
; FILING DATE: Herewith
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: 5202-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 498-8260
; TELEFAX: 617 876-5851
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-808-324-26

Query Match 71.4%; Score 1747; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.3e-141;
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 DLSHTPLRRQKYLFDVSMLSKDELVGAEELRLFRQAPSPAGPLHVQLFPCLSPLL 196
|||||